



Department of Economics

Essays on Time-Consistent Fiscal Policy

Joana Pereira

*Thesis submitted for assessment with a view to obtaining the degree of
Doctor of Economics of the European University Institute*

Florence
April 2009

EUROPEAN UNIVERSITY INSTITUTE
Department of Economics

Essays on Time-Consistent Fiscal Policy

Joana Pereira

Thesis submitted for assessment with a view to obtaining the degree of
Doctor of Economics of the European University Institute

Jury Members:

Prof. Robin Boadway, Queen's University, Ontario
Prof. Juan Carlos Conesa, Universitat Autònoma de Barcelona
Prof. Ramon Marimon, EUI
Prof. Salvador Ortigueira, EUI, supervisor

© 2009, Joana Pereira
No part of this thesis may be copied, reproduced or
transmitted without prior permission of the author

ACKNOWLEDGMENTS

This thesis has a single author but many participants. I am deeply grateful to my supervisors Salvador Ortigueira and Rick van der Ploeg for their guidance, encouragement and insightful discussions over the past four years. Their expertise, support and patience added considerably to my life as a PhD student at the European University Institute. I would also like to express my appreciation for their advice concerning the development of my career as an Economics researcher.

I thank Thomas Sargent for his comments on my early dissertation work and for hosting me at New York University's Economics Department during the Fall of 2006, where I strongly benefited from the interaction with faculty members and PhD students, namely the participants of his Macro Reading Group. I am also indebted to Morten Ravn, Bas Jacobs and my EUI colleagues not only for the helpful comments on my thesis, but also for their encouragement through the harder stages of my PhD experience.

A special thanks goes out to my family members, to Patrick and to my dear friends, without which support going through such a lonely task as writing a PhD dissertation would not have been possible.

Financial support from the Portuguese Foreign Affairs Ministry and the Fundação para a Ciência e Tecnologia is gratefully acknowledged.

CONTENTS

I	Introduction	3
II	Time-Consistent Fiscal Policy: Optimal Social Choices and the Role of Fiscal Constitutions	7
1	OPTIMAL EDUCATION SUBSIDIES: COMPARING THE SECOND BEST WITH A TIME-CONSISTENT REDISTRIBUTION POLICY	1
1.1	Introduction	1
1.2	The Baseline Economy	5
1.2.1	Individual Life-Cycle Problem	5
1.2.2	The Government Problem	7
1.3	Redistribution with Linear Policy Instruments	8
1.3.1	A Fictitious Representative Consumer	9
1.4	Optimal Fiscal Policy with Commitment	12
1.5	Optimal Fiscal Policy without Commitment	16
1.6	A Parameterized Economy	20
1.7	Conclusion	25
1.8	Appendix	25
1.9	References	33
2	TIME-CONSISTENT REDISTRIBUTION POLICY AND EDUCATION SUBSIDIES UNDER ALTERNATIVE FISCAL CONSTITUTIONS	34
2.1	Introduction	34
2.2	The Economy	37
2.3	Redistribution Policy with Complete Information	40
2.4	Optimal Nonlinear Redistribution Policy under Commitment	41
2.4.1	Nonlinear Policy with Limited Record Keeping: $T(s, Y) = T(Y)$. .	44
2.4.2	Time Inconsistency and the Ratchet Effect	46
2.5	Optimal Time-Consistent Nonlinear Redistribution Policy	48
2.5.1	Separation in Youth with Full Insurance in Adulthood	49
2.5.2	Pooling in Youth with Second-Best in Adulthood	53
2.5.3	Optimal Time-Consistent Nonlinear Redistribution Policy under Limited Record Keeping	55
2.6	Conclusion	59

<i>CONTENTS</i>	3
2.7 Appendix	60
2.8 References	67
3 MARKOV-PERFECT OPTIMAL FISCAL POLICY: THE CASE OF UNBALANCED BUDGETS	70
3.1 Introduction	70
3.2 The Model	73
3.2.1 Households	73
3.2.2 Firms	74
3.2.3 Government	74
3.2.4 Ramsey Optimal Fiscal Policy	75
3.2.5 Markov-Perfect Optimal Fiscal Policy	75
3.2.5.1 Steady-State Markov-Perfect Equilibrium	80
3.3 The Markov-Perfect Equilibrium in a Calibrated Economy	82
3.3.1 The Role of Debt in the Markov-Perfect Equilibrium	85
3.4 Numerical Approach	86
3.5 Conclusions	87
3.6 Appendix	88
3.7 References	96

Part I

Introduction

Governments are major economic players in all societies. The effect of fiscal policy - government spending, taxation and subsidies as well as public debt - on the aggregate economy and social welfare is of utmost importance and therefore subject to great controversy amongst economists. Indeed, the literature on optimal fiscal policy is impressively sizable and has tackled issues as varied as the provision of public goods, choice amongst different tax instruments in deterministic and stochastic settings, tax competition, credibility issues in dynamic economies and redistribution of income across heterogeneous agents. The traditional benchmark in all these normative analyses is the so called First-Best allocation. According to the First-Best theory of taxation, economies that are not subject to market frictions should not be distorted by government intervention in the optimum. Therefore, relative prices should be equalized to the respective marginal rates of transformation, which usually calls for using lump-sum transfers only. In the absence of first-best tools or in the presence of imperfections - such as externalities, informational asymmetries and missing markets that call for an optimal distortion of market prices - economists have resorted to the Second Best theory of taxation to prescribe optimal policy. If it is imperative to use distortionary tools, the Second Best allocation is always such that the implied total excess burden is effectively minimized. In dynamic economies, and in particular economies where investment decisions are relevant, such goal is usually achieved by reducing long run distortions at the expense of short-run taxation. The reason is simply that past investments are irreversible and thus a non-distortionary source of revenue, whilst current investments are mostly discouraged by a high taxation of future returns. Consequentially, the theory recommends a different policy rule for the initial and later periods. Notably, this is only feasible if private agents believe that in the future all policy makers will comply with the currently announced fiscal plan; that is, if governments can credibly commit to the intertemporal policy that is optimal as of date 0. That is arguably a very strong assumption, as even a benevolent government will often find it optimal to reset its fiscal programme as time unfolds.

This thesis consists of an analysis of different aspects of optimal fiscal policy in dynamic economies, with a special emphasis on the consequences of dropping the assumption of full commitment to future policies: the *Third Best* environment. As is now well known in the dynamic optimal taxation literature - namely after the seminal contributions by Kydland and Prescott (1977) and Fisher (1980) - the fact that governments choose their fiscal policy sequentially means that they cannot credibly make promises on future taxes that fall on investments made today. When the *future* arrives, previous choices are bygones and namely the previous investment decisions will have been translated into inelastically supplied factors of production [such as physical and human capital]. It will then be optimal to revise past

promises and tax heavily such inputs so as to lower the remaining distortions. As such, each government should set its policy as a best response to the expected, time-consistent, future policy rules. This generally represents also deviating from the second best solution for the first period.

The aim of my dissertation is to provide new insights on how should benevolent governments optimally set their fiscal plan in scenarios where policy is chosen sequentially [as it is done in real economies] and to contrast the obtained results both to what we observe in developed economies and to the Second Best. Furthermore, special emphasis will be put on the design of the fiscal constitutions and its role in alleviating the burden of policy discretion over social welfare. For that, I will focus on two particular examples of the time-inconsistency problem in intertemporal policy making, hoping that the lessons derived therefrom can be extended easily to qualitatively similar problems.

In the first part of the thesis - Chapters 1 and 2 - I extend the literature on optimal redistribution policy to a dynamic economy where skills are endogenously determined at the initial stage of the life cycle. Namely, I consider an economy where agents first choose how much to invest on their education - given their innate ability and initial economic resources - and only later enter the labor market and supply labor. The optimal redistribution plan for the entire life cycle of individuals is then characterized under the two mutually exclusive assumptions: either governments can credibly commit to future policies or they cannot. I find that in this economy an uncommitted utilitarian government setting taxes after individuals have entered the labor market will have an incentive to over-redistribute labor income [sets marginal tax rates too high] vis-a-vis the second best plan. In turn, that calls for first-period subsidies to education that are *regressive* - i.e., render the life-time tax system less progressive by benefitting relatively more those who will end up with higher income -, partially correcting for the over progressiveness of future labor taxes. In going from Chapter 1 to Chapter 2, I allow for greater flexibility in the fiscal instruments available to the government. Namely, Chapter 2 is a study of optimal non-linear redistribution policy, whilst Chapter 1 includes only linear taxation/subsidies. The novelty is that when a non-linear tax-schedule is allowed, the time-inconsistency problem is exacerbated by the so called Ratchet Effect, according to which full redistribution is possible after education decisions are sunk because the government faces no constraints to implementing the first best allocation. We show that tacit fiscal constitutions that are observed in real economies such as the independence of the tax code from past schooling choices are potentially welfare improving.

The third Chapter - which is joint work with Salvador Ortigueira - analyzes Markov-perfect optimal fiscal policy in a neoclassical economy with physical capital and public debt.

The possibility of running deficits/surpluses in this class of economies is the major contribution of the paper. Previous studies on Markov-perfect policy abstract from either public debt, by assuming a government's period-by-period balanced budget constraint, or from physical capital, assuming that labor is the only factor of production. We show that the steady-state Ramsey equilibrium (the second-best) is time-consistent. Moreover, for a standard parameterization of the economy we find a second, stable Markov-perfect equilibrium in which income taxes are positive and public debt is higher than 50% of GDP. These numbers are in line with those observed in most developed economies. The multiplicity arises from differences in expectations over future policy. A feature common to the two equilibria is that governments use public debt to reduce long-run tax distortions, as compared to the economy without debt. The requirement to run balanced budgets in every period is, thus, identified as a welfare-deterimental fiscal constitution.

Part II

Time-Consistent Fiscal Policy: Optimal Social Choices and the Role of Fiscal Constitutions

CHAPTER 1

OPTIMAL EDUCATION SUBSIDIES: COMPARING THE SECOND BEST WITH A TIME-CONSISTENT REDISTRIBUTION POLICY

1.1 Introduction

Most developed economies have witnessed an increase in pre-tax earnings' inequality during the last decades. Factors such as the rise of the skill premium and the increase in European unemployment rates, especially among less skilled individuals, have been identified as the major sources of such income disparities. As inequalities deepen, so does the demand for social insurance. In fact, the share of welfare spending on total government budgets has risen considerably and income tax bills are strongly progressive. These developments have revived an interest in assessing the welfare consequences of income inequality and on the uprising role of the welfare state as insurance provider.

The aim of this paper is to characterize optimal redistribution policy in scenarios where governments' commitment to future policies is restricted. In particular, our research will focus on the role of education subsidies as an instrument for redistribution. Education is understood as a principal determinant of labor productivity and hence of pre-tax wage inequality. Under this hypothesis, current schooling policies are bound to have an important influence over the future need for social insurance and optimal public choices. Moreover, such influence is mostly relevant when policy making is made sequentially, i.e., when each government maximizes social welfare given current state variables and regardless of what was chosen in the past. Because policy making during a big time span - specifically, from youth to adulthood of an individual - is likely to observe this feature in real economies, it is essential to understand the trade-offs involved in dynamic redistribution policy under lack of commitment.

The fact that income gaps appear empirically rooted in skill heterogeneity supports the modern approach to optimal redistribution problems. In a canonical Mirrlees(1971) economy, agents differ in a single parameter - usually called "ability" - that summarizes differences in labor productivity. This parameter, beyond the choice of individuals, is understood as the major source of lifetime uncertainty. In a world of risk-averse agents and inequality-averse

planners, redistribution policies are then desirably implemented. If there is perfect information on the relevant idiosyncrasies, that comes at no efficiency cost. Differentiated lump sum taxes are sufficient to implement a first best allocation with perfect insurance. Mirrlees seminal contribution is to provide a theory of second best taxation (with non-linear marginal taxes on labor income) in this context, given the obvious agency problems that arise by virtue of private information (non contractibility) on one's own productivity. Perfect insurance is not possible under asymmetric information because it would entail huge distortions on labor supply. But the government can still provide partial income redistribution, and it does so by optimally balancing the trade-off between equity gains and efficiency losses of the necessary distortionary tax-scheme.

In a recent paper, Werning (2007) explores the dynamics of income taxes in a neoclassical economy where the distribution of skills is subject to exogenous shocks over time. A planner chooses the time sequence of labor and capital taxes that maximizes social welfare, balancing the benefits and costs of the resulting income redistribution. The optimal plan is such that marginal labor taxes are set higher when the income gap is wider and vice-versa. Intuitively, the higher the level of inequality, the lower the weight of distortionary costs vis-a-vis equity gains. Thus, labor income taxes become more desirable. They are more effective in providing social insurance when the pre-tax wage gap is higher. Arguably, however, the distribution of skills and productivity gaps are not only changing exogenously over time and depend to a high extent on previous schooling choices. What role, then, for the education subsidies? How should marginal subsidies to education relate to income taxation? Observably, in most industrialized economies human capital investments prior to the entry in the labor market are heavily subsidized, often at marginal rates above those of labor income taxes. Is this optimal from a redistributive point of view? Does the time-consistency requirement affect the normative prescriptions of education policy?

To answer these questions, we develop a life-cycle model of heterogeneous agents, who differ both in ability to learn and in initial wealth. Young individuals choose how much consume and invest in their human capital, given their innate characteristics. They take as given current schooling policy and form expectations about future redistributive taxes. Later, when adults, they choose how much to consume and to work, given their acquired skills, accumulated savings/debt and the tax schedule chosen by the policy makers. The model is the simplest setting where one could explore the consequences of time-inconsistency in redistribution policy when labor productivity is *i)* endogenously determined, unlike in a traditional Mirrleesian setting and *ii)* depends - in the absence of social insurance - on a realistic set of different initial conditions: wealth and ability.

This project relates to a vast literature on optimal redistribution policy. Recent advances in dynamic public finance build on Mirrlees (1971) static framework by introducing an investment technology into the baseline problem. Albanesi and Sleet (2006), Kocherlakota (2005), Werning (2007) consider neoclassical capital accumulation economies to determine optimal dynamic taxation of capital and labor when skills are determined by idiosyncratic shocks. Bohacek and Kapicka (2008) discuss instead optimal education policies in a human capital accumulation with infinite horizon but a permanent shock to productivity. Their model generalizes the analysis made by Bovenberg, Jacobs (2005) for static economies, reaching similar conclusions on the role of education subsidies: if taxes are set optimally, education subsidies play a minor role as redistribution tools; rather, they are set essentially to restore efficiency on education investment.

The problem of discretion in a Mirrlees setup with endogenous skills was first studied by Boadway et al (1996). In their two-period economy, a time-inconsistency problem arises because governments may observe agent's productivity after education decisions are made. Full redistribution is then costless in the second period, but it causes huge disincentives to invest in human capital *ex ante*. A case for the high education subsidies that are observed in developed economies is then made. Notice, however, that in such an environment the lack of commitment is binding because redistributive taxes are optimally contingent on past schooling investment (albeit indirectly, through the resulting labor productivity). In fact, such policies are not at all in practice in real economies. Thus, we do not allow for this possibility in our model as we restrict attention to linear taxation instruments. Instead, the commitment problem will arise from a difference in initial conditions to the sequential utilitarian governments' problems: the distribution of skills and assets within the population is not necessarily the same at the beginning of lifetime and as upon entry in the labor market.

We find that when governments can credibly commit to their preferred fiscal plan in the first period, having a life-cycle structure is not much informative vis-a-vis considering a similar but static economy as the one studied by Bovenberg, Jacobs (2005, 2008): for a linear earnings function, optimal marginal education subsidies are exactly equal to future marginal labor taxes when the elasticity of complementarity between schooling and innate ability in building up human capital is equal to one. Since labor tax rates reduce proportionately the final return on schooling, such policy corresponds to a zero net subsidy on human capital investments. When there is complementarity (substitutability) between ability and schooling in accumulating skills, though, the social planner will prefer to tax (subsidize) education in net terms for redistribution purposes. What is new in our setup is the conclusion that initial wealth differences are optimally redistributed away at the beginning of lifetime and that

savings shall not be distorted at the optimum. Both results are consistent with Werning (2007). In contrast, uncommitted governments first redistribute against differences in initial conditions - ability and wealth - but later are only concerned with the resulting distribution of acquired skills and assets. Previous sources of inequality and endogenous schooling decisions are disregarded. Thus, insofar as schooling boosts skill heterogeneity, optimal labor taxation is more distortionary *ex post* than *ex ante*. Likewise, *ex post* differences in savings would be completely eliminated if the government could choose any savings tax (this also means that debts would eventually be bailed out). From the perspective of the first period policy maker, this redistribution choice balances the "wrong" trade-off between equity and efficiency. Thus, education policy is used not only to restore efficiency in human capital investment (in the same way as in the commitment economy), but to restore the relevant trade-off between equity and efficiency from the beginning of life-time.

Education subsidies are redistributive insofar as they complement labor taxes when these are chosen suboptimally. For a standard parameterization of our economy, we find this to uphold an optimal marginal education subsidy above the future labor tax. Intuitively, the tax set upon entry in the labor market is over-redistributive vis-a-vis the second best plan (that of a committed government). Then, the best response of the first period government is a compensation through an education policy that benefits future high income earners relatively more. In fact, Bohacek and Kapicka(2008) also support the case for education subsidies as a tool for redistribution when income labor taxes are set exogenously. Their environment is considerable more complex than the one in this paper but a parallel may be done to their results. When labor taxes are set optimally, redistribution is carried out essentially through those taxes. Subsidies to education are introduced mostly to restore efficiency¹. By contrast, when labor taxes are set exogenously, education subsidies compensate for the resulting suboptimality. Indeed, the main conclusion on education policy within our setup is comparable. Only the source of inefficiency is not an exogenous assumption on the level of the labor tax rate, but the time inconsistency inherent in the dynamic redistribution policy of utilitarian governments.

In the next section we set up our baseline economy, identifying the problem of the individuals and their optimization conditions as well as the scope of government intervention in the economy. Section 1.3 describes problem of the social planner with access to a general set of linear fiscal instruments, and lays out a primal approach to that problem, adapted from Werning (2007). Sections 1.4 and 1.5 compare the optimal redistribution policy under

¹In their paper the resulting subsidy rate is quantitatively small (3%) because costs of education are already deductible from the tax bill. The remaining effect owes to the type-dependency of education costs (as foregone earnings, which depend on type) and the effect that subsidies have on incentive compatibility.

full commitment and under discretion for a general specification of individual utility and technology. We then parameterize the economy in Section 1.6 and provide a numerical example on the differences in optimal income taxes and education subsidies under each case. Finally, Section 1.7 concludes.

1.2 The Baseline Economy

1.2.1 Individual Life-Cycle Problem

Our economy is populated by a continuum of individuals taking economic decisions in two periods of their life cycle. In period 1, young individuals decide how to split their available resources between consumption, investment in human capital and savings. They may differ in cognitive ability $\theta \in \Theta$ and inherited resources $a^0 \in \Lambda$. When they become older, in period 2, individuals choose how much to work and consume, given an idiosyncratic labor productivity which depends on previous schooling decisions. Thus, agents in this economy are divided into a finite number of types $i \in I = \Theta \times \Lambda$ of relative size π_i , which are defined as $i \equiv (\theta, a^0)$. We will abuse notation and denote θ_i and a_i^0 the specific characteristics of an individual of (global) type i . Unitary labor productivity is then given by a function $\psi(\theta_i, s_i)$, where s_i denotes the investment in schooling previously chosen by agent i . The skill production function $\psi(\cdot)$ observes positive diminishing returns to each of its arguments. Furthermore, there is complementarity between schooling and ability - $\psi_{s\theta} > 0, \forall s$ - and $\lim_{s \rightarrow 0} \psi_s = +\infty$.

All individual characteristics are assumed to be private information, and in particular the cognitive ability θ . Only its distribution across individuals is common knowledge. This is a standard assumption in the Mirrlees taxation framework. In practice, it implies that individual abilities are not taxable per se, but only as determinants of total labor income.

Agents maximize lifetime utility, assumed to be separable in time and between goods and leisure:

$$U^i = u(c_i^1) + \beta [u(c_i^2) - v(n_i)], \quad (1.1)$$

with $u_c(\cdot) > 0$, $u_{cc}(\cdot) < 0$, $v_n > 0$, $v_{nn} < 0$. The assumption of separability is made for simplicity, and does not affect the main insights of the paper. All agents share the same preferences over the consumption of normal goods and leisure. However, individuals with different skill levels at the beginning of the second period derive different disutility from generating the same labor income. This follows from a production technology which is assumed to be linear in labor (the only input) and perfect competition in the production sector, so that all workers are paid their marginal productivity - $y_i = n_i \psi(\theta_i, s_i)$.

Education is costly in terms of money and is paid for in the first period. The cost may include tuition fees, purchases of goods and services directly related to education and/or foregone earnings during youth². Thus, the sequential problem faced by the individual can be described as follows. In the first period, agent i solves

$$\begin{aligned} V_i^1(a_i^0, \theta_i, T_i^1, T_i^2) &= \max_{c_i^1, s_i, a_i^1} u(c_i^1) + \beta V_i^2(\theta_i, s_i, a_i^1, T_i^2) \\ \text{s.t.} \quad c_i^1 + s_i + a_i^1 &\leq a_i^0 - T_i^1 \\ a_i^0 &\text{ given,} \end{aligned} \quad (1.2)$$

where V_i^1 and V_i^2 are the indirect utilities of a type i agent when young and when old, respectively, as functions of individual state variables. The latter is given by the second period problem

$$\begin{aligned} V_i^2(\theta_i, s_i, a_i^1, T_i^2) &= \max_{c_i^2, y_i} \left[u(c_i^2) - v\left(\frac{y_i}{\psi(\theta_i, s_i)}\right) \right] \\ \text{s.t.} \quad c_i^2 &\leq a_i^1(1+r) + y_i - T_i^2. \end{aligned} \quad (1.3)$$

$1+r$ is the exogenously given interest rate and T_i^1 and T_i^2 are, respectively, the net tax bills of agent i in periods 1 and 2. We define their functional form bellow. Individuals neglect the effect of their individual choices on aggregate variables such as the overall distribution of income/skills, meaning that T_i^1 and T_i^2 are taken as given.

First order conditions for the problem an agent of type i are, then:

$$MRS_{s_i, c_i^1} = \frac{\beta v_n(n_i) \frac{\psi_s(\theta_i, s_i)}{\psi(\theta_i, s_i)} n_i}{u_c(c_i^1)} = 1 - \frac{\partial T_i^1}{\partial s_i} \quad (1.4)$$

$$MRS_{c_i^1, c_i^2} = \frac{u_c(c_i^1)}{\beta u_c(c_i^2)} = 1 + r - \frac{\partial T_i^1}{\partial a_i^1} \quad (1.5)$$

$$MRS_{y_i, c_i^2} = \frac{v_n(n_i)}{u_c(c_i^2) \psi(\theta_i, s_i)} = 1 - \frac{\partial T_i^2}{\partial y_i} \quad (1.6)$$

$$c_i^1 + s_i + a_i^1 = a_i^0 - T_i^1 \quad c_i^2 = a_i^1(1+r) + y_i - T_i^2$$

At this point, it matters the distinction between a scenario where the policy maker announces the tax schedule for the whole life time and commits to it and one in which she lacks credibility and therefore agents have to form expectations on what will be the tax

²It would be straightforward to include labor supply in the first period and model explicitly these foregone earnings, as in Bovenberg and Jacobs (2005, 2008), Kapicka (2006) or Bohacek and Kapicka (2007). How results are affected depends on the possibility or not of directly subsidizing/taxing time devoted to learning and on the specific way in which education time enters in the human capital production function $\psi(\cdot)$, -interaction between time and resources devoted to education and between education time and innate ability. However, as long as labor taxes can be age-dependent abstracting from foregone earnings does not impact our major conclusions about time-consistency in policy making.

schedule for each period. Comparing redistribution policy in each of the two scenarios is indeed the contribution of this paper. In the former, we may think of T_i^2 as "known" to all agents from birth. In the latter, agents form expectations on its future value which we denote \tilde{T}_i^2 . In a rational expectations equilibrium, all agents correctly foresee the future tax schedule and therefore necessarily share the same beliefs about future policy.

Marginal taxes are zero when the economy is either left to a *laissez faire* or if government levies / transfers are lump-sum. In those cases, the allocation of resources within the economy is fully efficient as all marginal rates of substitution are equated to the respective relative market prices.

Finally, notice that schooling investment is distorted not only by the net relative price of education but also the net relative price of leisure in the second period. From (1.4)-(1.6), optimality for human capital investment may be rewritten:

$$MRS_{s_i, c_i^1} = \frac{\beta u_c(c_i^2) \psi_s(\theta_i, s_i) n_i \left(1 - \frac{\partial T_i^2}{\partial y_i}\right)}{u_c(c_i^1)} = 1 - \frac{\partial T_i^1}{\partial s_i} \quad (1.7)$$

Hence, human capital investment is subsidized (taxed) in net terms when $\frac{1 - \partial T_i^1 / \partial s_i}{1 - \partial T_i^2 / \partial y_i}$ is strictly lower (bigger) than 1. When free borrowing/saving is allowed, (1.7) together with (1.5) define a non-arbitrage condition for investment in assets versus schooling. Therefore, savings' taxes will also affect human capital accumulation.

1.2.2 The Government Problem

We consider a benevolent government choosing allocations - i.e., setting fiscal policy - so as to maximize an Utilitarian social welfare function.

$$V^{soc} = \sum_{i \in I} U^i(c_i^1, c_i^2, n_i) \pi_i \quad (1.8)$$

This government is conditioned by the set of available fiscal instruments and the degree of commitment to future policies. In this paper, we focus on a simple case where all fiscal instruments are linear, meaning that only aggregate income, schooling investment and savings are observable. Under such fiscal constitution, we explore the optimal lifetime redistribution policy, contrasting the second best normative results to the - perhaps more realistic - time-consistent solution. The extension to a non-linear redistribution policy is discussed in Pereira (2008).

For simplicity, we abstract from any revenue requirements other than those created by the distributional concern. Finally, governments may borrow/save in international capital markets any amount B , bearing the same interest rate as private agents: r .

Thus, welfare maximization is subject to the aggregate resource constraints (capital letters denote aggregates over all types):

$$C^1 + S + A^1 \leq A^0 + B \quad (1.9a)$$

$$C^2 \leq (A^1 + B)(1 + r) + Y \quad (1.9b)$$

A Walras law applies, meaning that when all individual budget constraints hold with equality, the budget of the government also holds with equality.

1.3 Redistribution with Linear Policy Instruments

Take the simplest case where all fiscal instruments are linear but the set of available taxes/subsidies is otherwise complete. Governments may tax initial asset holdings, labor income and returns from savings according to the flat rates τ_{a^0} , τ_y and τ_{a^1} , respectively. The tax bill may include also a lump-sum component $-T^t$ (T is to be interpreted as a transfer) in each period, which means that fiscal policy is not automatically distortionary. Finally, the cost of investing in human capital is subsidized at the rate x , independent of the schooling level or initial wealth. In sum, individual budget constraints are, $\forall i$,

$$c_i^1 + s_i(1 - x) + a_i^1 = a_i^0(1 - \tau_{a^0}) + T^1, \quad (1.10)$$

$$c_i^2 = a_i^1[1 + r(1 - \tau_{a^1})] + y_i(1 - \tau_y) + T^2. \quad (1.11)$$

In this scenario, governments can *independently* set each of the implied margins in (1.4)-(1.6). The only feature that distinguishes it from a pure Mirrleesian choice is that optimal marginal rates of substitution are necessarily equalized across agents: $MRS_{s_i, c_i^1} = (1 - x)$, $MRS_{c_i^1, c_i^2} = [1 + r(1 - \tau_{a^1})]$ (or $= [1 + r(1 - \bar{\tau}_{a^1})]$ if the government can credibly commit to future policy) and $MRS_{y_i, c_i^2} = (1 - \tau_y)$ for all i . As such, the implied division of aggregates $\{C^1, C^2, A^1, S, Y\}$ within the population is necessarily efficient. Any distortions caused by fiscal policy will be restricted to the determination of those four variables. That is because linearity in the fiscal instruments is equivalent to observability of aggregates only. Therefore, if the government chooses to distort the economy because of equity concerns, it can only do so over what it "observes" (what may condition the tax schedule on).

One may, then, formalize the government's problem as a choice over aggregate allocations instead of over tax rates. The approach is proposed by Werning (2007) in a dynamic economy with idiosyncratic shocks to labor productivity and aggregate uncertainty. We apply it here to our deterministic life-cycle economy with some necessary modifications due to i) the direct

impact that choices in the first period have on second period's disutility of working, and ii) eventual lack of commitment by the initial policy maker to future redistribution policy.

1.3.1 A Fictitious Representative Consumer

Following Werning (2007), we first establish an exact correspondence between the set of first order conditions from all individual problems - describing a competitive equilibrium of our economy, for each tax policy - and those of the optimization program we set up below. This will allow us to write individual allocations and relative prices as functions of aggregate variables.

Consider first the competitive equilibrium arising in the second period, after differences in skills are permanently set. In this period, all agents face the same marginal tax rate on income. Thus, allocation of effective labor (y_i) and consumption (c_i^2) across the different types is undistorted, given the total Y and C^2 . That is, for each competitive equilibrium (each tax rate) there exist positive market weights $\{\varphi_i^2\}$, normalized so that $\sum_{i \in I} \varphi_i^2 \pi_i = 1$, such that individual allocations can be obtained from the following programme:

$$\begin{aligned}
 U^{2,m}(C^2, Y^2, \{\varphi_i^2\}) = \max_{\{c_i^2, y_i\}} \sum_{i \in I} \varphi_i^2 \left[u(c_i^2) - v\left(\frac{y_i}{\psi(\theta_i, s_i)}\right) \right] \pi_i \\
 \text{s.t.} \\
 \sum_{i \in I} c_i^2 \pi_i = C^2 \quad \sum_{i \in I} y_i \pi_i = Y
 \end{aligned} \tag{1.12}$$

This defines an optimal *market* division of C^2 and Y , given the existing distribution of $\{\psi(\theta_i, s_i)\}_{i \in I}$. The solution consists of a set of individual allocations $\{c_i^2, y_i\}_{i \in I}$, of which each element is a function of $(C^2, Y, \{\varphi_i^2\})$.³ The weights represent the importance that the allocation of an agent of type i has in the indirect utility of the "representative consumer" - the market, in all - in the second period. They strongly hinge on the degree of income redistribution in the economy and initial inequality, as is discussed in Section 1.4. To see that the solution to (1.12) corresponds to the outcome of competitive equilibrium, take first order conditions and make use of the envelope theorem to obtain:

$$\begin{aligned}
 \varphi_i^2 u_c(c_i^2) &= U_{C^2}^{2,m} \\
 -\varphi_i^2 \frac{v_n(n_i)}{\psi(\theta_i, s_i)} &= U_Y^{2,m}
 \end{aligned} \tag{1.13}$$

³We will not explicitly include these as arguments of individual allocations when stating the social planner's problem due to space constraints. There is, however, no possible ambiguity in reading the problem. Wherever we refer to government optimization, individual variables are to be automatically understood as functions of aggregates and market weights.

A fictitious representative consumer with utility function $U^{2,m}(C^2, Y, \{\varphi_i^2\})$ and budget constraint $C^2 \leq Y(1 - \tau_y) - T^2$ would, then, face the same relative price of consumption versus leisure as all other agents in this economy:

$$\frac{-U_Y^{2,m}}{U_{C^2}^{2,m}} = (1 - \tau_y). \quad (1.14)$$

Thus,

$$\frac{\varphi_i^2 \frac{v_n(n_i)}{\psi(\theta_i, s_i)}}{\varphi_i^2 u_c(c_i^2)} = \frac{-U_Y^{2,m}}{U_{C^2}^{2,m}} = MRS_{y_i, c_i^2}$$

so that we obtain the same set of first order conditions as in the decentralized equilibrium. Finally, one may substitute the tax rate in the individual budget constraint (1.11) by the expression in (1.14), and reach the set of implementability constraints:

$$U_{C^2}^{2,m} c_i^2 + U_Y^{2,m} y_i = U_{C^2}^{2,m} [1 + r(1 - \tau_{a^1}) + T^2], \forall i \in I \quad (1.15)$$

which may be written exclusively in terms of $(C^2, Y, \{\varphi_i^2\}, T^2)$ and the initial distribution of skills and assets.

A similar exercise is now done for the youth allocations. The *market* optimization in the first period involves not only intra-period felicity derived from c_i^1 , for each i , but also the effect of s_i and a_i^1 on the continuation value V_i^2 . In sum, the market programme for the first period is

$$\begin{aligned} U^{1,m}(C^1, S, \{\varphi_i^1\}; \{c_i^2\}, \{y_i\}) &= \max_{c_i^1, s_i, a_i^1} \sum_{i \in I} \varphi_i^1 [u(c_i^1) + \beta V_i^2(s_i, a_i^1)] \pi_i \\ &\quad s.t. \\ \sum_{i \in I} c_i^1 \pi_i &= C^1 \quad \sum_{i \in I} s_i \pi_i = S \quad \sum_{i \in I} a_i^1 \pi_i = A^1 \end{aligned} \quad (1.16)$$

$\underline{V}_i^2(s_i, a_i^1) \equiv V_i^2(\theta_i, s_i, a_i^1, T_i^2)$, with $V_i^2(\cdot)$ as defined in (1.3). As of this period, $\{c_i^2\}, \{y_i\}$ are either exactly known, as functions of $(\bar{C}^2, \bar{Y}, \{\bar{\varphi}_i^2\})$ - aggregates to which the government may commit to (as determined by its tax policy) - or they are functions of the foreseen, time-consistent, $(\tilde{C}^2, \tilde{Y}, \{\tilde{\varphi}_i^2\})$. Problem (1.16) is otherwise unaffected by the commitment assumption. Therefore, one may simply write the market's indirect utility as $U^{1,m}(C^1, S, \{\varphi_i^1\}; C_i^2, Y, \{\varphi_i^2\})$ and then interpret the variables accordingly. Taking the same steps as above and making use of the envelope theorem in (1.3) to find $\frac{dV_i^2(s_i, a_i^1)}{ds_i}$, it is straightforward to derive

$$\frac{U_S^{1,m}}{U_{C^1}^{1,m}} = \frac{\varphi_i^1 v_n(n_i) \psi_s(\theta_i, s_i) n_i}{\varphi_i^1 u_c(c_i^1) \psi(\theta_i, s_i)} = 1 - x \quad (1.17)$$

and, by substituting in (1.10),

$$U_{C^1}^{1,m}(c_i^1 + a_i^1) + U_S^{1,m}s_i = U_{C^1}^{1,m}[a_i^0(1 - \tau_{a^0}) + T^1], \forall i \in I \quad (1.18)$$

To completely establish the correspondence between the competitive equilibrium in period 1 and problem (1.16), however, one needs to consider in addition the individual choice of savings for all types. We show in the Appendix that the latter implies that a decentralization of the socially optimal allocation is only possible when the market weights $\{\varphi_i^1\}$ and $\{\varphi_i^2\}$ coincide:

Lemma 1.1 *In an economy with perfect capital markets where individual allocations are obtained as the solution to (1.16)-(1.12), the market weights in (1.16) necessarily observe the condition*

$$\varphi_i^1 = \tilde{\varphi}_i^2, \forall i \in I, \quad (1.19)$$

with $\tilde{\varphi}_i^2 = \bar{\varphi}_i^2$ if there is perfect policy commitment and $\tilde{\varphi}_i^2$ equal to φ_i^2 in equilibrium.

In general, (1.19) implies that policy makers can only choose $\{\varphi_i^2\}$ directly. In order to implement the optimal aggregate allocation under perfect capital markets the government cannot generate a different weight of individual i 's utility in the *market* total utility (the utility of the representative consumer). Since the decentralization is made through a system of linear prices, all individuals face the same (expected) intertemporal price of consumption. Would individual i have a higher weight than individual j when young but not when old, this would mean that at least one of the agents could improve her lifetime utility by changing a_i^1 and *smoothing utility*. Therefore, we could not be in the presence of a competitive equilibrium. Notice that Lemma 1 relies on the time-separability of preferences but not on separability between consumption and leisure.

The following proposition, adapting Proposition 1 in Werning (2007) to our model, sums up the main content of this subsection:

Proposition 1.2 *Given the initial distribution of net assets, $\{(1 - \tau_{a^0})a_i^0\}_{i \in I}$, an aggregate allocation (C^1, C^2, A^1, S, Y) can be supported by a competitive equilibrium in our economy if it is feasible and there exist market weights $\{\varphi_i^1\}$, $\{\varphi_i^2\}$ and lump-sum transfers T^1, T^2 such that (1.19) holds and implementability constraints (1.15)-(1.18) are verified for all $i \in I$. Individual allocations of young agents can then be obtained as the solution to problem (1.16), whereas those of old agents derive from (1.12).*

1.4 Optimal Fiscal Policy with Commitment

We are now in the position of characterizing the second best social problem. Since the planner may credibly commit to future taxes, she indeed chooses at date zero all lifetime aggregates and second period market weights. Thus, making use of condition

$$U_{C^1}^{1,m} / \beta U_{C^2}^{2,m} = [1 + r(1 - \tau_{a^1})] \quad (1.20)$$

from the "fictitious representative consumer" problem, we may rewrite (1.15) and (1.18) as the set of present value implementability constraints

$$U_{C^1}^{1,m} c_i^1 + U_S^{1,m} s_i + \beta U_{C^2}^{2,m} c_i^2 + \beta U_Y^{2,m} y_i = U_{C^1}^{1,m} [a_i^0 (1 - \tau_{a^0}) + T], \forall i \in I, \quad (1.21)$$

where T is the present value of all lump-sum transfers. The committed government maximizes (1.8) subject to aggregate feasibility (also in present value) and market implementability - conditions (1.21) and (1.19) - by choosing $(C^1, C^2, S, Y, \{\varphi_i^2\})$, the lump sum transfer T and τ_{a^0} (which is not distortionary but can be redistributive). The implicit taxes that are indirectly being chosen follow (1.14), (1.17) and (1.20).

Two comments are in order. First, the choice of market weights $\{\varphi_i^2\}$ is associated with the embedded degree of inequality aversion in the social welfare function. For our parameterization of the latter, social inequality aversion pertains to the concavity of the utility function (1.1).⁴ The higher the social value of equity, the more similar these weights will be, meaning that allocations tend to be more similar too (see, e.g., (1.13)). In turn, achieving similar consumption levels across differently productive individuals requires setting relatively high proportional tax rates and thus having lower aggregate income and consumption. If these inefficiencies would not follow from redistribution of income (e.g., if factors were supplied inelastically), the optimal market weights would be chosen exactly equal to the Pareto weights in social welfare - see last footnote -, which in (1.8) are set to 1 for all types. Second, the lower is type heterogeneity, the weaker the insurance motive to achieve a high social welfare. In the limit, when all agents are equal, optimal weights can be set to 1 at no cost, as determined in the market, and the government raises lump-sum taxes only if there are eventual revenue requirements (which we have ruled out in this paper).

The Lagrangian for the government problem, with $\mu_i \pi_i$ the multiplier of i 's imple-

⁴For a more general social welfare function $\sum_{i \in I} \lambda_i U^i \pi_i$, it would also reflect the structure of the Pareto weights $\{\lambda_i\}_{i \in I}$.

mentability constraint (1.21) and (1.19) appropriately substituted for is given by:

$$\begin{aligned}
& \sum_{i \in I} [U^i(c_i^1, c_i^2, n_i) + \mu_i (U_{C^1}^{1,m} c_i^1 + \beta U_{C^2}^{2,m} c_i^2 + U_S^{1,m} s_i + \beta U_Y^{2,m} y_i)] \pi_i - \\
& - U_{C^1}^{1,m} \sum_{i \in I} \mu_i [a_i^0 (1 - \tau_{a^0}) + T] \pi_i - \lambda \left[A^0 - C^1 - K + \frac{Y - C^2}{1 + r} \right] \\
= & W(C^1, C^2, S, Y, \{\varphi_i^2\}, \{\mu_i\}) - \\
& - U_{C^1}^{1,m} \sum_{i \in I} \mu_i [a_i^0 (1 - \tau_{a^0}) + T] \pi_i - \lambda \left[A^0 - C^1 - K + \frac{Y - C^2}{1 + r} \right]
\end{aligned} \tag{1.22}$$

Notice that μ_i represents the marginal social benefit of resources being assigned specifically to type i agents; i.e., net of the positive effect through the correspondent increase in total resources. Only when taxes/subsidies are non distortionary will this net effect be zero.

By optimizing with respect to T and τ_{a^0} , one reaches the conditions

$$\sum_i \mu_i \pi_i = 0 \tag{1.23}$$

$$\sum_i \mu_i a_i^0 \pi_i = 0 \text{ or } \tau_{a^0} = 1 \tag{1.24}$$

Together, they imply that the second term of the Lagrangian is not binding; it will disappear from (1.22). Each of them understates fairly intuitive results. (1.23) implies that the marginal social benefit of increasing resources for a particular type is on average equal to that of increasing economy-wide resources. Hence, the net effect has mean zero. In turn, (1.24) implies that the government resorts to a full redistribution of observable initial resources. As is well understood in the literature, optimal redistribution policy balances a non-trivial trade-off between equity benefits and efficiency costs of distortionary taxation. The latter can, however, be substantially reduced if consumption smoothing within the population is attained mostly by virtue of a simple reshuffle of initial resources. τ_{a^0} is a purely efficient tax instrument, just like T , but the former serves the redistributive purpose whilst the latter does not. When a_i^0 is constant, that is unprofitable, and τ_{a^0} is indeterminate.⁵

Given (1.23) and (1.24), we now describe the necessary conditions for optimality in the

⁵In practice, (1.24) implies that initial wealth heterogeneity plays little role in the optimal redistribution plan. Would we have $a_i^0 = a^0, \forall i$, the remaining results of the paper would be unaffected. However, we chose to maintain it here for the sake of generality and because we are currently working on an extension of the model to unobservable $\{a_i^0\}_{i \in I}$, in which the heterogeneity assumption no longer is inconsequential.

choice of aggregate variables. In our deterministic economy they take a very simple form:

$$\left\{ \begin{array}{l} W_{C^1}/W_{C^2} = 1 + r \\ W_S/W_{C^1} = 1 \\ -W_Y/W_{C^2} = 1 \end{array} \right. \xRightarrow{\text{(by (1.14), (1.17), (1.20))}} \left\{ \begin{array}{l} \frac{U_{C^1}^{1,m}/\beta U_{C^2}^{2,m}}{W_{C^1}/W_{C^2}} = \frac{1+r(1-\bar{\tau}_{a^1})}{1+r} \\ \frac{U_S^{1,m}/U_{C^1}^{1,m}}{W_S/W_{C^1}} = (1 - \bar{x}) \\ \frac{U_Y^{2,m}/U_{C^2}^{2,m}}{W_Y/W_{C^2}} = (1 - \bar{\tau}_y) \end{array} \right. \quad (1.25)$$

Upper bars are used to emphasize that these are tax choices the government pre-commits to.

$W(\cdot)$ is to be interpreted as the *pseudo* social welfare function- social utility net of binding implementability constraints -, which the government maximizes subject to the intertemporal resource constraint. The system on left hand side of (1.25) is the social parallel to (1.4)-(1.6), the difference being that governments care about before tax prices and have a different objective function. Thus, distortions are imposed to private choices whenever the social marginal rate of substitution between two of the goods - (aggregate) consumption, leisure and schooling - is different from that implied in the *market* (the fictitious representative consumer). For instance, individual supply of labor is optimally distorted because the government valorizes less higher aggregate consumption (as opposed to a more equal consumption) than the market, where there is no intrinsic preference for equity. The social marginal rate of substitution of consumption for leisure in the second period will then be as low relative to that of the market, as more inelastic the labor supply. The other margins have similar interpretations.⁶

Education Policy under Commitment

Werning (2007) presented similar expressions for the dynamic taxation of capital and labor⁷. What is new and of special interest to us in this paper is how $(1 - x)$ is determined; namely, how it relates to other distortions imposed on the economy and which role it plays in the redistribution plan. We can write it as

$$\begin{aligned} (1 - \bar{x}) &= (1 - \bar{\tau}_y) \frac{W_Y/W_S}{U_Y^{2,m}/U_S^{1,m}} \frac{W_{C^1}/W_{C^2}}{U_{C^1}^{1,m}/U_{C^2}^{2,m}} \\ &= \frac{1 + r}{1 + r(1 - \bar{\tau}_{a^1})} (1 - \bar{\tau}_y) \left[\frac{W_Y/W_S}{\beta U_Y^{2,m}/U_S^{1,m}} \right] \end{aligned} \quad (1.26)$$

Ignore for the moment the last term. Subsidies to education will, on the one hand, exactly compensate for the distortions induced by taxation of labor income - $(1 - \bar{\tau}_y)$ -, as

⁶It is worth noting that expressions in (1.25) are not explicit formulae for taxes/subsidies, as the ratios involved in it depend yet on endogenous variables.

⁷See also Atkeson, Chari and Kehoe (1999) for an homogeneous agents - infinite horizon economy.

it is through labor income that the investment on schooling meets a return. Such result is widely documented in Kapicka(2006), Bohacek and Kapicka(2008) and Bovenberg and Jacobs (2005). Additionally, \bar{x} decreases when the optimal savings tax is higher, as investment on education becomes a more attractive form of transferring resources across time. This was to be expected by looking at the non-arbitrage condition (1.7). Since the government can always resort to lump-sum taxes/transfers in the first period to cover revenue requirements, it can fully compensate for the inefficiency costs caused by these two indirect distortions over the investment on human capital by simply introducing the inverse wedge.

Whether or not schooling will be encouraged in net terms depends finally on the size of the ratio $\frac{W_Y/W_S}{\beta U_Y^{2,m}/U_S^{1,m}}$. To interpret it, take the individual optimization problem. The marginal rate of substitution between investment in education and leisure is given by $n_i \psi_s(\theta_i, s_i)$. It tells us by how much will labor income, and thus utility derived from consumption and leisure, rise at the margin after investing in education. In terms of the aggregate allocation, such margin will only differ from the planner's and the *market's* points of view if schooling benefits high and low types at different rates. Namely, if the elasticity of complementarity between schooling and ability is higher (lower) than 1, the rate of increase in productivity grows with innate ability and the government will value education less than the market. In other words, the social marginal rate of substitution will be lower (higher) than that of the market - $\frac{W_Y/W_S}{\beta U_Y^{2,m}/U_S^{1,m}} < (>) 1$ - and education will be net taxed (subsidized) for redistributive reasons. The insight is discussed extensively in Bovenberg and Jacobs (2008)⁸.

In sum, this subsection extends previous results by Bovenberg, Jacobs (2005, 2008) and Reis (2005) who study a similar redistribution problem, with optimal education policy and income taxation, in static economies. Because the government has full commitment and there is no uncertainty in the model, redistribution policy set in the beginning of lifetime simply insures *ex ante* against initial heterogeneity and it does so by fixing at date zero the whole life path for individual allocations. The fact that they are spread in two periods is inconsequential for education policy.

Time Inconsistency

Assuming that governments commit to such redistribution policy is, arguably, very restrictive. From the moment when agents take their major schooling decisions, during their youth, to the span of their participation in the labor market there is a considerable time interval. Suppose the utilitarian government is allowed to re-optimize at the beginning of the

⁸The authors also discuss the impact on optimal education subsidies of a non-unitary elasticity of complementarity between labor effort and schooling in the earnings function. For our specification - $y = \psi(\theta, s)n$ - this elasticity is equal to 1.

second period. It will by then be faced with a single period *Mirrleesian* economy where the population is characterized by a known distribution unverifiable skills $\{\psi_i\}_{i \in I}$, not necessarily equivalent to that of the pre-education innate abilities $\{\theta_i\}_{i \in I}$. The scope of heterogeneity may have waxed or waned. Furthermore, the social objective function will be different; it is now simply the average utility of adult individuals. Thus, the redistribution problem changes as time unfolds.

Preferred taxes as of the beginning of the second period generally differ from what was optimal one period before, a constraint that the first period policy maker faces in the absence of long term commitment. In other words, a time-inconsistency problem arises which is particularly tied to the assumption of utilitarianism: what matters at each stage of policy revision are the established sources of heterogeneity - e.g. $\{\psi_i\}_{i \in I}$ - and how best to *redistribute* them away; not where do they come from.

What is then a time-consistent redistribution policy in this model? We turn to the no-commitment environment in the next section.

1.5 Optimal Fiscal Policy without Commitment

Three features distinguish the problem faced by the government (re)optimizing at the beginning the second period from that of a committed planner who chose ex ante all lifetime allocations. One, education decisions are now sunk, which is equivalent to saying that the elasticity of labor supply is perceived lower than in the first period. There is a strong complementarity between early investment in human capital and hours of work during adulthood. Thus, the opportunity cost of leisure increases in the second period because productivity ψ has increased with education. Two, inequality is now deeper because $\psi_{\theta_s} > 0$. As such, redistribution policy is more prone to sacrifice efficiency in favor of closer consumption levels across types. Implementability constraints will, therefore bind differently for all types. Three, decisions on savings are also sunk at beginning of the second period. If there is heterogeneity in a_i^1 , and by a similar argument to the one following equation (1.23), τ_{a^1} now becomes a non distortionary tool for redistribution.

If we allow the government to freely choose τ_{a^1} at date 1, it is clear that the optimal social choice at that point will be $\tau_{a^1} = 1$. Consequently, every agent will choose to borrow infinitely, as she will then be completely bailed out in the future. Therefore, in order to have an interesting comparison with the second best solution in our stylized economy we need to impose either a limit to the much agents can borrow or the much governments may tax. A borrowing limit would always be trivially hit by all agents, however, changing the nature of the social programme: even in the absence of any equity concerns it will be

optimal to distort the price of schooling and labor (see e.g. Jacobs (2002)) so as to facilitate consumption smoothing across the different individuals. Hence, we will impose instead a limit to the much governments can tax savings and we set that limit to zero without loss of generality (they would otherwise choose whatever limit we might impose).

Because we assume that the consecutive governments are benevolent, there is no conflict in what concerns second period *equity goals*. If it was not for the implementability constraints, both governments would have all marginal utilities equalized across agents. But they disagree on the appropriate balance against efficiency losses. We argue in this section that subsidies to education then have a redistributive role, in that they not only compensate for the indirect distortions caused by labor income taxes but will be used by the first government to effectively redistribute against the initial shock θ and not differences in ψ .

Time-Consistent Redistribution Policy: Adulthood

The time-consistent problem of the government may be solved backwards. In the last period, the economy inherits the previously determined distribution of labor productivity and assets together with public assets/liabilities B . Social welfare is simply given by

$$V^{2,soc}(C^2, Y, \{\varphi_i^2\}) = \max \sum_{i \in I} [u(c_i^2) - v(n_i)] \pi_i$$

which the government maximizes over $(C^2, Y, \{\varphi_i^2\})$ and T^2 . All distortionary effects on past decisions are disregarded, as they are irrelevant to the current period problem. Furthermore, the government is constrained by aggregate feasibility and the set of implementability conditions (1.15). As such, the Lagrangian for this problem is

$$\begin{aligned} & \sum_{i \in I} [U^{2,i}(c_i^2, n_i) + \mu_i^2 (U_{C^2}^{2,m} c_i^2 + U_Y^{2,m} y_i)] \pi_i - \\ & - U_{C^2}^{2,m} \sum_{i \in I} \mu_i^2 [a_i^1 (1+r) + T^2] \pi_i - \lambda_2 [(A^1 + B)(1+r) + Y - C^2] \\ & = W^2(C^2, Y, \{\varphi_i^2\}, \{\mu_i^2\})|_{\{s_i\}_{i \in I}} - U_{C^2}^{2,m} \sum_{i \in I} \mu_i^2 [a_i^1 (1+r) + T^2] \pi_i \\ & - \lambda_2 [(A^1 + B)(1+r) + Y - C^2] \\ & = \mathcal{W}^2(C^2, Y, \{\varphi_i^2\}, \{\mu_i^2\}, T^2)|_{\{a_{i,s_i}^1\}_{i \in I}} - \lambda_2 [(A^1 + B)(1+r) + Y - C^2] \end{aligned} \quad (1.27)$$

The choice of T^2 is dictated by a similar condition to (1.23), turning the Lagrangian independent of T^2 . Thus, parallel to Section 1.3, $-\mathcal{W}_Y^2/\mathcal{W}_{C^2}^2 = 1$ holds in the social optimum, implying

$$\frac{\mathcal{W}_{C^2}^2/\mathcal{W}_Y^2}{U_{C^2}^{2,m}/U_Y^{2,m}} = (1 - \hat{\tau}_y) \quad (1.28)$$

The interpretation of this ratio is the same as under commitment. The higher the gap between marginal rates of substitution of consumption for leisure to the utilitarian government and to market, the bigger is τ_y . However, as is made clear in the example below (see also last subsection of the Appendix), the resulting policy rule for labor income tax is different for two reasons. One, labor productivity levels are now fixed, so the government no longer considers the complementarity between Y and S (that is, ignores the effect of Y on the first period implementability constraints). Consequentially, the optimal social balance between equity - more equal c_i^2 - and efficiency - higher Y and C^2 - changes, as the efficiency costs of taxation decrease. In the *market*, instead, commitment is not an issue because there is no intrinsic taste for redistributing resources. As a result, the optimal tax rate $\hat{\tau}_y$ tends to be higher than $\bar{\tau}_y$ so as to bring the equilibrium closer to a more equitable outcome. Two, unlike $\bar{\tau}_y$, $\hat{\tau}_y$ is going to be affected by the heterogeneity in asset holdings at the beginning of the second period because there is now a missing fiscal instrument to tax savings. The fiscal authority would like to bail out debtors and expropriate savers. Not being able to do it, $\hat{\tau}_y$ is bound to depend on the correlation between individual asset holdings and their skills. Intuitively, one would expect that a high correlation between the two would render $\hat{\tau}_y$ less effective as wealthier individuals have a lower marginal utility of income and therefore less willingness to work for the same ψ . If, on the contrary, the most productive individuals arrive at period 2 with lower savings - e.g. because they have borrowed more to invest on education - we could expect optimal $\hat{\tau}_y$ to increase even more vis-a-vis $\bar{\tau}_y$. Nevertheless, the expressions for $\bar{\tau}_y$ and $\hat{\tau}_y$ are sufficiently involving that a clear cut comparison of the two is impossible⁹.

Time-Consistent Redistribution Policy: Youth

In this period, governments regard the final distribution of skills and savings as the outcome of endogenous decisions, therefore taking it into account when setting fiscal policy. They are, however, restricted by the discretion of the next policy maker, who is also an utilitarian, and the inability to redistribute initial resources ($\tau_{a^0} = 0$) and set market weights $\{\varphi_i^1\}_{i \in I}$ freely. The social problem is thus confined to choosing (C^1, A^1, S) , B and T^1 , so as to maximize the average lifetime utility:

$$V^{1,soc} = \max_{\{C^1, S, A^1, B, T^1\}} \sum_{i \in I} u(c_i^1) \pi_i + \beta \tilde{V}^{2,soc}(S, A^1)$$

subject to the implementability constraints (1.18), the conditions on the market weights (1.19), and the future policy rules - future redistribution policy as a function of the inherited

⁹See last section of the Appendix.

distribution of $\{\psi_i, a_i^1\}_{i \in I}$ and $A^1 + B$. Finally, the redistribution programme faces an additional implementability constraint associated with the inability to set the future tax rate on savings. Hence, the correspondent Lagrangian function is

$$\begin{aligned}
& \sum_{i \in I} \{u(c_i^1) + \mu_i^1 [U_{C^1}^{1,m}(c_i^1 + a_i^1) + U_S^{1,m} s_i]\} \pi_i + \beta \tilde{V}^{2,soc}(S, A^1) \\
& - \sum_{i \in I} \mu_i^1 U_{C^1}^{1,m}(a_i^0 + T^1) \pi_i - \lambda_1 [A^0 - C^1 - A^1 - S] \\
& = W^1(C^1, A^1, S, \{\varphi_i^1\}, \{\mu_i^1\}) - U_{C^1}^{1,m} \sum_{i \in I} \mu_i^1 (a_i^0 + T^1) \pi_i - \lambda_1 [A^0 - C^1 - A^1 - S] \\
& = \mathcal{W}^1(C^1, A^1, S, \{\varphi_i^1\}, \{\mu_i^1\}, T^1) \big|_{\{a_i^0\}_{i \in I}} - \lambda_1 [A^0 - C^1 - A^1 - S]
\end{aligned} \tag{1.29}$$

When choosing education policy, the government takes the expectations of private agents on future policy as given. Whereas the existence of perfect private capital markets impose condition (1.19), the Euler equation (1.5) with zero taxes completely pins down $\{a_i^1\}_{i \in I}$ as a function of current and expected redistribution policy, as part of the set of necessary conditions for a competitive equilibrium. In the second best programme it was still possible to indirectly set $\{a_i^1\}_{i \in I}$ through the choice of τ_{a^1} .

In the social optimum, one obtains again that $\mathcal{W}_{C^1}^1/\mathcal{W}_S^1 = \mathcal{W}_{C^1}^1/\mathcal{W}_A^1 = \mathcal{W}_{C^1}^1/\mathcal{W}_B^1 = 1$. Combining it with (1.17) and rearranging terms, the optimal education subsidy rate becomes:

$$\begin{aligned}
(1 - \hat{x}) &= (1 - \tilde{\tau}_y) \frac{\beta \tilde{\mathcal{W}}_Y^2 / \mathcal{W}_S^1}{\beta \tilde{U}_Y^{2,m} / U_S^{1,m}} \frac{\mathcal{W}_{C^1}^1 / \beta \tilde{\mathcal{W}}_{C^2}}{U_{C^1}^{1,m} / \beta \tilde{U}_{C^2}^{2,m}} \\
&= (1 - \tilde{\tau}_y) \tilde{\xi} \frac{\tilde{\mathcal{W}}_Y^2 / \mathcal{W}_S^1}{\tilde{U}_Y^{2,m} / U_S^{1,m}}.
\end{aligned} \tag{1.30}$$

The elements in (1.30) are the exact parallel to those in (1.26), the same way that income taxes have similar expressions under commitment and discretion. However, $\tilde{\mathcal{W}}_Y^2$ is not part of the first government programme, meaning it is distinct from \mathcal{W}_Y^2 in (1.26). The first policy maker does internalize the fact that $\tilde{\mathcal{W}}^2$ ultimately depends on S , but this derivative is embedded in \mathcal{W}_S^1 as she does not directly control $\tilde{\mathcal{W}}^2$ (see last section of the Appendix). For the same reason, $\frac{\mathcal{W}_{C^1}^1 / \tilde{\mathcal{W}}_{C^2}^2}{U_{C^1}^{1,m} / \tilde{U}_{C^2}^{2,m}} = \tilde{\xi}$ no longer corresponds to a fully controllable fiscal price, but to a social intertemporal discount which depends on the initial choice of aggregate consumption and expectations about future policy.

In contrast, from the *market* perspective the only thing that has changed is that future variables are now foreseen, rather than known. The derivatives of $U^{t,m}|_{t=1,2}$ with respect to S and Y - are thus the same as in (1.26) after expectations realize. Therefore, the relative

marginal rates of substitution - to the market versus the utilitarian society - are bound to change.

One would ultimately like to have a clear contrast between \bar{x} and \hat{x} . Once more, given the complexity of the terms involved in (1.26) and (1.30) - evident from the Appendix - that is not possible. This is why we now turn to a familiar parameterization of the model that allows for a neater comparison of second best with time-consistent policy.

1.6 A Parameterized Economy

In order to illustrate the differences between commitment and discretion in the economy analyzed above, we particularize the results for an economy which functional forms allow for a near closed form solution. We assume an utility function of the form (1.1), with $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$ and $v(n) = \alpha \frac{n^\gamma}{\gamma}$, implying a Frisch elasticity of labor supply of $\frac{1}{\gamma-1}$. Skill technology is given by $\psi(\theta, s) = \theta^\eta s^\rho$. We assume all parameters to be strictly positive and $\rho < 1$. In sum,

$$U^i = \frac{(c_i^1)^{1-\sigma}}{1-\sigma} + \beta \left[\frac{(c_i^2)^{1-\sigma}}{1-\sigma} - \alpha \frac{n^\gamma}{\gamma} \right]$$

$$\psi(\theta, s) = \theta^\eta s^\rho \tag{1.31}$$

$$0 < \sigma \leq 1, 0 < \beta < 1, \gamma > 1, \alpha > 0.$$

In this case, the allocation problem of the *pseudo* representative consumer has a simple solution, with all individual variables determined as shares of their aggregate correspondents:

$$\begin{aligned} c_i^1 &= \omega_c^i C^1 & c_i^2 &= \omega_c^i C^2 \\ y_i &= \omega_y^i Y & s_i &= \omega_s^i S. \end{aligned}$$

Derivations and expressions for ω_c^i , ω_s^i and ω_y^i may be found in the Appendix. These shares are functions of the parameters of the model and the market weights, which are taken as given by individuals. Notice that the individual allocations do not depend on A^1 . Both in the commitment and no commitment solutions, the $\{a_i^1\}_{i \in I}$ may be obtained as functions of the other policy variables. As the same happens with the public savings B , the optimality conditions for these two aggregates are actually the same - there is no increased benefit in using public debt/savings when agents can perfectly smooth consumption in the private market.

Redistribution Policy Under Commitment

Proposition 1.3 *For the parameterization (1.31) of the economy laid out in Section 1.2, and when governments may credibly commit to future fiscal policy, optimal taxes on savings are set to zero - $\bar{\tau}_{a^1} = 0$ - and the optimal labor income tax and subsidies to education are given by:*

$$\bar{\tau}_y = \bar{x} = 1 - \frac{E \left[\frac{1}{\varphi_i^2} \right] + \text{cov} \left(\omega_c^i, \frac{1}{\varphi_i^2} + \mu_i (1 - \sigma) \right)}{E \left[\frac{1}{\varphi_i^2} \right] + \text{cov} \left(\omega_y^i, \frac{1}{\varphi_i^2} + \mu_i \gamma \right) - \gamma \rho \text{cov} \left(\omega_s^i, \mu_i \right)} = 1 - \frac{\sum_i \omega_c^i \left[\frac{1}{\varphi_i^2} + \mu_i (1 - \sigma) \right] \pi_i}{\sum_i \omega_y^i \left[\frac{1}{\varphi_i^2} + \mu_i \gamma (1 - \rho) \right] \pi_i}$$

with μ_i defined as in (1.22).

Proof. See Appendix. ■

Committed governments dispose of savings taxes and use instead labor taxes to redistribute income. Distorting the intertemporal price of consumption brings no additional insurance against initial inequality. Neither does it help alleviate the distortionary effects of labor tax rates due to the separability between consumption and leisure. Thus, for this parameterization of the utility function, there is no intertemporal distortion.

The simple relation between the labor tax rate and education subsidies is a consequence of the technology for skill accumulation having an unitary elasticity of substitution between innate ability and schooling. The ratio $\frac{W_Y/W_S}{\beta U_Y^{2,m}/U_S^{1,m}}$ in (1.26) is then equal to 1. Furthermore, the Cobb-Douglas technology for skill accumulation implies $\omega_s^i = \omega_y^i$, simplifying the expression for $\bar{\tau}_y$.

$\bar{\tau}_y$ balances the trade-off between the benefits of redistributing labor income and its distortionary costs. As mentioned before, at the beginning of lifetime one shall think of labor supply as being composed of time not enjoying leisure and resources not allocated to initial consumption, both of which having the purpose of increasing labor income. The elasticity of this "composite" labor supply is bigger than the regular one and thus the effective redistribution costs for the same γ are higher. This justifies the $(1 - \rho)$ term in the denominator, which wouldn't be there if skills would be permanently equal to θ .

When $\text{cov} \left(\omega_c^i, \frac{1}{\varphi_i^2} + \mu_i (1 - \sigma) \right)$ decreases with risk aversion σ , welfare derived from insurance rises and $\bar{\tau}_y$ is optimally higher. On the other hand, for $\text{cov} \left(\omega_y^i, \frac{1}{\varphi_i^2} + \mu_i \gamma \right)$ increasing in γ , a big elasticity of labor supply - low γ - implies high social costs of taxation and hence a lower optimal $\bar{\tau}_y$. Likewise, the impact of a change in returns to schooling in the optimal labor taxation depends on how $\gamma \rho \text{cov} \left(\omega_s^i, \mu_i^1 \right)$ changes with ρ . A bigger ρ generates, in equilibrium, higher aggregate investment on schooling, S , and thus higher Y for the same total hours worked in the economy. However, the ex-ante elasticity of labor supply (the elasticity

of the "composite" labor supply) also increases, counteracting the first effect. Whether the optimal tax increases/decreases when ρ is higher is, thus, a numerical question.

Time Consistent Redistribution Policy

If second period governments have discretion over fiscal policy, they set redistribution policy after schooling choices were made, and hence labor supply is no longer regarded as the "composite" of foregone leisure and first period consumption. Rather it is simply foregone leisure, as in a classical Mirrlees environment. Therefore, the relevant elasticity is the pure Frisch elasticity of labor supply, determined by γ . Moreover, the government is hampered by the impossibility to redistribute away differences in a_i^1 . The optimal choice of labor taxes is thus qualitatively different from the commitment case. Proposition 1.4 states the optimal formulas for $\hat{\tau}_y$ and \hat{x} :

Proposition 1.4 *For the parameterization (1.31) of the economy laid out in Section 1.2, when governments cannot credibly commit to future fiscal policy and may not levy taxes (bail out) on savings (borrowing), the optimal labor income tax and subsidies to education are given by:*

$$\begin{aligned}\hat{\tau}_y &= 1 - \frac{\sum_i \omega_c^i \left[\frac{1}{\varphi_i^2} + \mu_i^2 (1 - \sigma) \right] \pi_i + \frac{\sigma}{C^2} \sum_i \mu_i^2 a_i^1 (1 + r) \pi_i}{\sum_i \omega_y^i \left[\frac{1}{\varphi_i^2} + \mu_i^2 \gamma \right] \pi_i} = \\ &= 1 - \frac{E \left[\frac{1}{\varphi_i^2} \right] + \text{cov} \left(\omega_c^i, \frac{1}{\varphi_i^2} + \mu_i^2 (1 - \sigma) \right) + \frac{\sigma(1+r)}{C^2} \text{cov} (\mu_i^2, a_i^1)}{E \left[\frac{1}{\varphi_i^2} \right] + \text{cov} \left(\omega_y^i, \frac{1}{\varphi_i^2} + \mu_i^2 \gamma \right)}\end{aligned}$$

and

$$1 - \hat{x} = (1 - \tilde{\tau}_y) \frac{E \left[\frac{1}{\tilde{\varphi}_i^2} \right] + \text{cov} \left(\tilde{\omega}_y^i, \frac{1}{\tilde{\varphi}_i^2} + \tilde{\mu}_i^2 \gamma \right)}{E \left[\frac{1}{\tilde{\varphi}_i^2} \right] + \text{cov} \left(\tilde{\omega}_y^i, \frac{1}{\tilde{\varphi}_i^2} + \tilde{\mu}_i^2 \gamma \right) - \gamma \rho \text{cov} \left(\omega_s^i, \frac{1}{1+\beta} \mu_i^1 + \frac{\beta}{1+\beta} \tilde{\mu}_i^2 \right)}$$

where μ_i^1 and μ_i^2 are the multipliers of the implementability constraints for agent i in (1.29) and (1.27).

Proof. See Appendix. ■

Comparing with the expressions in Proposition 1.3, the covariances of ω_c^i , ω_s^i and ω_y^i with the multipliers μ_i^1 and μ_i^2 will certainly differ because the latter do not coincide and equilibrium market weights change. Yet, for the same covariance between the share of labor

income and μ_i^2 , it is now less costly to tax labor. The denominator is then higher, suggesting that $\hat{\tau}_y$ will tend to be higher than $\bar{\tau}_y$. Also, $\hat{\tau}_y$ is now affected by the heterogeneity in savings. Suppose that agents who are more skilled and therefore generate a higher share of total earnings have borrowed more / saved less in the past to finance their studies. Then, if $\text{cov}(\omega_y^i, \mu_i^2) > 0$ - as is the case in the example below -, we must have $\text{cov}(a_i^1, \mu_i^2) < 0$, contributing to the rise in $\hat{\tau}_y$. When the opposite happens (both covariances are positive), for instance due to a high correlation between θ_i and a_i^0 , the optimal tax on labor decreases because of a wealth effect over labor supply decisions (labor supply becomes more elastic).

Subsidies to education, as in the commitment case, compensate for the indirect distortions that labor taxes impose. Nonetheless, their net value depends on how ω_s^i co-varies now with μ_i^1 and $\tilde{\mu}_i^2$. Since for our parameterization $\omega_s^i = \omega_y^i$, $\hat{x} > \hat{\tau}_y$ if and only if $\text{cov}\left(\tilde{\omega}_y^i, \frac{1}{1+\beta}\mu_i^1 + \frac{\beta}{1+\beta}\tilde{\mu}_i^2\right) < 0$, which requires $\text{cov}(\tilde{\omega}_y^i, \mu_i^1) < 0$. Thus, the optimal education policy relies on how $\{\mu_i^1\}_{i \in I}$ and $\{\tilde{\mu}_i^2\}_{i \in I}$ relate. Note that these are multipliers to the same present value implementability constraint, which is perceived differently by the two governments, not only because of schooling and savings ex-ante endogeneity and ex-post exogeneity but because the fiscal instruments available to each governor are not the same. Together, $\{\tilde{\mu}_i^2\}_{i \in I}$ measure the excess burden of taxation in the second period and $\{\mu_i^1\}_{i \in I}$ have a similar interpretation for schooling subsidization. Only when the multipliers are inversely correlated will net education subsidies ($\hat{x} > \tilde{\tau}_y$) be socially desirable.

Calibration

We now calibrate the parameters in the utility function, skill technology and initial heterogeneity. Given the stylized form of our economy, we do not attempt to exactly match properties of the data. Instead, we choose a meaningful calibration that allows us to get a quantitative notion of the direction of our results. Hence, in what concerns the utility function we set $\sigma = 1$ (logarithmic utility in consumption), $\gamma = 5$ (implying a Frisch elasticity of labor supply equal to 0.25, as in Bohacek, Kapicka (2008)) and normalize $\alpha = 1$. The skill technology will be linear in cognitive ability θ , i.e., $\eta = 1$. The curvature with respect to the level of education - ρ - is set to different levels for comparison purposes, ranging from 0:3 to 0:7: Finally, $\beta = (1 + r)^{-1} = (1 + 0.04)^{-1}$.

The economy is populated by two types only: H and L . The initial abilities θ_H and θ_L are such that their average is fixed to 1 and we make the ratio θ_H/θ_L grow to see the response of optimal tax rates and education subsidies to a rise in inequality as of the beginning of lifetime. As we derived above, heterogeneity in initial resources under the second best is always fully redistributed, and is thus inconsequential for the other policy variables. As

such, to keep results comparable under the commitment and no-commitment solutions (as we artificially imposed zero taxes on wealth in the latter), we assume in this example that all agents start with the same $a_i^0 = A^0$. The precise level of A^0 is qualitatively unimportant and we thus set it to zero.

The Figures in the Appendix display optimal income taxes and education subsidies under the two commitment scenarios. In the second best environment, taxes grow with initial inequality in θ , but tend to an upper bound well below 100%. Education subsidies naturally match labor income taxes, as derived above for our particular parameterization. Importantly, whilst at low levels of inequality a higher return on schooling in the skill technology allows for higher taxes and thus more redistribution, when inequality is too high, the excess burden becomes excessive. As the "composite" elasticity of labor supply increases with ρ , optimal tax rates at high inequality levels are lower when ρ is higher.

The present value of lump sum transfers is positive at all considered (ρ, θ_H) , regardless of government credibility. That means both that tax collection in the second period more than covers first period public expenditure in education and that income taxes are progressive. Overall, the fiscal plan is the more progressive the higher is τ_y and the lower is x .

As expected, labor taxation is higher under lack of commitment because the endogeneity of skills is not internalized. Moreover, the effect of an increase in ρ on the optimal tax level is stronger for the initial levels of inequality and does not reverse at high levels, precisely because the internalized elasticity of labor supply is simply the Frisch elasticity. In this scenario, it is actually the case that for $\rho = 0.7$ taxes approach quickly the 100% level ($\hat{\tau}_y$ never actually reaches it, but the numerical difference is negligible). At that stage aggregate consumption is virtually zero. In all, from a second best perspective, there is too much redistribution when taxes are chosen upon entry in the labor market and therefore welfare is lower.

Finally, education is always subsidized in net terms for our calibration. This means that governments at the beginning of lifetime consider the future tax rate as over redistributive, and thus use a higher-than-efficient subsidy to education that benefits relatively more future high income earners to compensate for such over-redistribution. Given that the model does not allow for a fully closed form solution, we cannot go beyond numerical simulations to assert the generality of such proposition. However, for a reasonable range of parameter values, we consistently found this pattern.

Both Bovenberg, Jacobs (2005) and Bohacek, Kapicka (2008) show that second best redistribution policy has education subsidies/taxes playing a redistributive role when labor

taxes are set exogenously¹⁰, and hence suboptimally. The interpretation here is equivalent, only the source of suboptimality is not an arbitrary value assumed for τ_y but the lack of commitment. Therefore x is also chosen to offset the unavoidable inefficiencies associated with an ill-chosen labor tax.

1.7 Conclusion

Previous literature on optimal redistribution policy with endogenous skills has found that education subsidies should be used in the fiscal plan mostly to restore efficiency in human capital investment. In this paper, we characterize optimal redistribution policy in a life-cycle economy where agents differ in their innate ability and initial wealth. In the considered setup, as expected, pure second-best policies (under government commitment) basically resemble what has been obtained in literature for static economies, namely in what concerns education policy.

Our main contribution is to drop the assumption of government's commitment to future policy in this context. We find that the lack of commitment is restrictive and curbs the extent to which labor income taxes insure optimally against initial inequality. In particular, we showed with a standard parameterization of the economy that the fiscal authorities "over-redistribute" labor income, as the second period idiosyncrasies are wider than first period ones and the elasticity of education choices is disregarded. Thus, governments at the beginning of lifetime see future labor tax policies as not adequately balancing the ex-ante trade-off between equity and efficiency. Education policy then plays a redistributive role insofar as it does not simply restore efficiency in investment decisions but partially undoes the excessive redistribution carried out in the second period.

1.8 Appendix

Proof of Lemma 1:

The first order conditions to problems (1.12) and (1.16) with respect to the individual consumption allocations and savings are given by

$$\begin{cases} \varphi_i^2 U_{c_i^2}^{2,i} = \lambda_{C^2} = U_{C^2}^{2,m}; \forall i \\ \varphi_i^1 U_{c_i^1}^{1,i} = \lambda_{C^1} = U_{C^y}^{1,m}; \forall i \\ \varphi_i^1 U_{a_i^1}^{1,i} = \lambda_A = U_A^{1,m}; \forall i \end{cases}$$

¹⁰Bovenberg and Jacobs (2005) consider a scenario where labor income cannot be taxed, whereas Bohacek and Kapicka (2008) set the labor tax rate to a flat 40%.

Take the problem of an agent i , who chooses allocations for the first period, forming expectations on future policy. From the first period problem (1.2) we have, by construction $U_{a_i^1}^{1,i} = \beta \frac{d\tilde{V}_i^2(\cdot)}{da_i^1}$ and, by an envelope argument applied to the second period problem (1.3), $\frac{d\tilde{V}_i^2(\cdot)}{da_i^1} = U_{c_i^2}^{2,i} [1 + r(1 - \tau_{a^1})]$. The margin between consumption today and consumption tomorrow, as a function of expectations on future government choice, is then:

$$\frac{U_{c_i^1}^{1,i}}{\beta \tilde{U}_{c_i^2}^{2,i}} = 1 + r(1 - \tilde{\tau}_{a^1}).$$

Since

$$1 + r(1 - \tilde{\tau}_{a^1}) = \frac{U_{C^1}^{1,m}}{\beta \tilde{U}_{C^2}^{2,m}} = \frac{\varphi_i^1 U_{c_i^1}^{1,i}}{\tilde{\varphi}_i^2 \beta \tilde{U}_{c_i^2}^{2,i}}, \forall i \in I,$$

by the solution to (1.12) and (1.16) and the optimization of the "fictitious representative consumer", $\varphi_i^1 / \tilde{\varphi}_i^2$ is necessarily a constant and because market weights are normalized to have an average of 1, $\varphi_i^1 = \tilde{\varphi}_i^2, \forall i$. ■

The fictitious representative consumer problem for parameterization (1.31):

In the last period, for a given distribution of skills $\{\psi(\theta_i, s_i)\}_{i \in I}$ and policy $(C^2, Y, \{\varphi_i^2\}_{i \in I})$, and by Proposition 1, individual allocations are obtained from:

$$\begin{aligned} U^{2,m}(C^2, Y, \{\varphi_i^2\}) &= \max_{\{c_i^2, y_i\}} \sum_i \varphi_i^2 \left[\frac{(c_i^2)^{1-\sigma}}{1-\sigma} - \alpha \frac{\left(\frac{y_i}{\psi(\theta_i, s_i)}\right)^\gamma}{\gamma} \right] \pi_i \\ &\quad s.t. \\ \sum_i c_i^2 \pi_i &= C^2 \quad \sum_i y_i \pi_i = Y, \end{aligned}$$

which solution is easily reached after taking first order conditions:

$$\left\{ \begin{array}{l} (c_i^2) : \quad \varphi_i^2 (c_i^2)^{-\sigma} = \lambda_{C^2} = U_{C^2}^{2,m} \\ (y_i) : \quad -\frac{\varphi_i^2}{\psi(\theta_i, s_i)} \left(\frac{y_i}{\psi(\theta_i, s_i)}\right)^{\gamma-1} = \lambda_Y = U_Y^{2,m} \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} c_i^2 = \frac{(\varphi_i^2)^{1/\sigma}}{\sum_i (\varphi_i^2)^{1/\sigma} \pi_i} C^2 = \omega_c^{2,i} C^2, \\ y_i = \frac{(\varphi_i^2)^{\frac{1}{1-\gamma}} \psi(\theta_i, s_i)^{\frac{\gamma}{1-\gamma}}}{\sum_i (\varphi_i^2)^{\frac{1}{1-\gamma}} \psi(\theta_i, s_i)^{\frac{\gamma}{1-\gamma}} \pi_i} Y = \omega_y^i(\{s_i\}) Y, \\ n_i = \frac{\omega_y^i(\{s_i\})}{\psi(\theta_i, s_i)} Y = \omega_n^i(\{s_i\}) Y \end{array} \right.$$

where λ_{C^2} and λ_Y are the multipliers to the aggregation constraints. Substituting the solution in the expression for $U^{2,m}(\cdot)$, we obtain the indirect utility of the *market* (i.e., fictitious representative consumer):

$$U^{2,m}(C^2, Y, \{\varphi_i^2\}) = \Phi_u^{2,m} u(C^2) - \Phi_v^{2,m}(\{s_i\}) v(Y),$$

with

$$\begin{aligned}\Phi_u^{2,m} &= \sum_i \varphi_i^2 (\omega_c^{2,i})^{1-\sigma} \pi_i = \left[\sum_i (\varphi_i^2)^{1/\sigma} \pi_i \right]^\sigma, \\ \Phi_v^{2,m}(\{s_i\}) &= \sum_i \varphi_i^2 (\omega_n^{2,i}(\{s_i\}))^\gamma \pi_i = \left\{ \sum_i \left[(\varphi_i^2)^{\frac{1}{1-\gamma}} \psi(\theta_i, s_i)^{\frac{\gamma}{1-\gamma}} \right] \pi_i \right\}^{1-\gamma}.\end{aligned}$$

Likewise, for a given distribution of abilities $\{\theta_i\}_{i \in I}$, current policy (C^1, S, A^1) and foreseen aggregates - functions of the expected tax schedule - the first period allocation solves:

$$\begin{aligned}U^{1,m}(C^1, S, A^1) &= \max_{\{c_i^1, s_i, a_i^1\}} \sum_i \varphi_i^1 \left[\frac{(c_i^1)^{1-\sigma}}{1-\sigma} + \beta \tilde{V}^{2,i}(s_i, a_i^1) \right] \pi_i \\ &\quad s.t. \\ \sum_i c_i^1 \pi_i &= C^2 \quad \sum_i s_i \pi_i = S \quad \sum_i a_i^1 \pi_i = A^1.\end{aligned}$$

The first order conditions for this problem are:

$$\begin{cases} (c_i^1) : & \varphi_i^1 (c_i^1)^{-\sigma} = \lambda_{C^1} = U_{C^1}^{1,m} \\ (s_i) : & \varphi_i^1 \beta \tilde{V}_{s_i}^{2,i} = \lambda_S = U_S^{1,m} \\ (a_i^1) : & \varphi_i^1 \beta \tilde{V}_{a_i^1}^{2,i} = \lambda_A \end{cases},$$

where λ_{C^1} , λ_S and λ_A are the multipliers to the aggregation constraints. By the envelope theorem applied to the individual problem, $V_{s_i}^{2,i} = v_n(n_i) n_i \frac{\psi_s(\theta_i, s_i)}{\psi(\theta_i, s_i)}$ and $V_{a_i^1}^{2,i} = (c_i^2)^{-\sigma} (1+r)$. Then, using the solution to the second period decentralization derived above:

$$\begin{aligned}\varphi_i^1 &= \tilde{\varphi}_i^1, \\ c_i^1 &= \frac{(\tilde{\varphi}_i^2)^{1/\sigma}}{\sum_i (\tilde{\varphi}_i^2)^{1/\sigma} \pi_i} C^1 = \tilde{\omega}_c^{2,i} C^1, \\ s_i &= \frac{\left[(\tilde{\varphi}_i^2)^{\frac{1}{1-\gamma}} (\theta_i^\eta)^{\frac{\gamma}{1-\gamma}} \right]^{\frac{1}{1-\rho \frac{\gamma}{1-\gamma}}}}{\sum_i \left[(\tilde{\varphi}_i^2)^{\frac{1}{1-\gamma}} (\theta_i^\eta)^{\frac{\gamma}{1-\gamma}} \right]^{\frac{1}{1-\rho \frac{\gamma}{1-\gamma}}} \pi_i} S = \omega_s^i S.\end{aligned}$$

In equilibrium, expectations are correct and we have $\tilde{\omega}_c^{2,i} = \omega_c^{2,i} = \omega_c^i$. Furthermore, after substituting s_i in $\omega_y^i(\{s_i\})$, we conclude that $\omega_y^i = \omega_s^i$ in equilibrium. The expression for a_i^1 is unnecessary to solve the social problem when there is government commitment. For the time-consistent solution, it may be derived from the implementability condition $u_c(c_i^1) = \beta(1+r)u_c(\tilde{c}_i^2) \Leftrightarrow c_i^1 = \tilde{c}_i^2$, making use of (1.18) and the individual budget constraints in

(1.3) as expected by the agents. This defines a_i^1 as a function of $\tilde{\tau}_y$, current policy (C^1, S, T^1) and individual characteristics: a_i^1, θ_i . We may finally replace the equilibrium values in the expression for $U^{1,m}(\cdot)$ to define it as a function of current policy:

$$U^{1,m}(C^1, S, T^1) = \Phi_u^{1,m} u(C^1) + \beta \sum_i \tilde{\varphi}_i^2 \left[\tilde{V}^{2,i}(a_i^1(C^1, S, T^1), s_i(S)) \right] \pi_i,$$

where $\Phi_u^{1,m} = \tilde{\Phi}_u^{2,m}$. ■

Proof of Proposition 1.3:

Under full government commitment, all variables are known up-front. Then,

$$\begin{aligned} U^{1,m} &= \Phi_u^{1,m} u(C^1) + \beta U^{2,m}, \\ U^{2,m} &= \left[\Phi_u^{2,m} u(C^2) - \bar{\Phi}_v^{2,m} \frac{v(Y)}{S^{\gamma\rho}} \right], \end{aligned}$$

where $\bar{\Phi}_v^{2,m} = \left\{ \sum_i \left[(\varphi_i^2)^{\frac{1}{1-\gamma}} (\theta_i^\eta (\omega_s^i)^\rho)^{\frac{\gamma}{1-\gamma}} \right] \pi_i \right\}^{1-\gamma}$ is equal to $\Phi_v^{2,m}(\{s_i\})$ after the $\{s_i\}_{i \in I}$ are appropriately substituted for. Using the solution to the fictitious representative consumer, and after tedious algebra, the expression for $W(\cdot)$ as defined in (1.22) becomes

$$W(C^1, C^2, S, Y; \{\varphi_i^2\}) = \Phi_u^W u(C^1) + \beta \left[\Phi_u^W u(C^2) - \Phi_v^W \frac{v(Y)}{S^{\gamma\rho}} \right],$$

with:

$$\begin{aligned} \Phi_u^W &= \Phi_u^{2,m} \sum_i \omega_c^i \left[\frac{1}{\varphi_i^2} + \mu_i (1 - \sigma) \right] \pi_i, \\ \Phi_v^W &= \bar{\Phi}_v^{2,m} \sum_i \bar{\omega}_y^i \left[\frac{1}{\varphi_i^2} + \mu_i \gamma (1 - \rho) \right] \pi_i; \quad \bar{\omega}_y^i = \omega_y^i(\{s_i\}) \text{ after the } \{s_i\}_{i \in I} \text{ are substituted.} \end{aligned}$$

Taking derivatives of $U^{1,m}$ with respect to C^1 and S , of $U^{2,m}$ w.r.t. C^2 and Y and of $W(\cdot)$ w.r.t all its arguments and replacing in (1.25), we obtain the expressions in Proposition 2 and the zero tax on savings. ■

Proof of Proposition 1.3:

The steps to derive the time consistent tax and education subsidy are similar to those taken to derive Proposition 1.3. A major difference is that the second period government chooses $\{\varphi_i^2\}$ without replacing $\{s_i\}_{i \in I}$ by their equilibrium expressions (which do depend on the weights at time 2, as shown above). Indeed, $\mathcal{W}^2(\cdot)$ becomes, after replacing c_i^2 and y_i by $\omega_c^i C^2$ and $\omega_y^i(\{s_i\}) Y$ respectively and remembering there are no taxes on savings,

$$\mathcal{W}^2(C^2, Y; \{\varphi_i^2\}, \{\mu_i\}) = \Phi_u^{2,W} u(C^2) - \Phi_v^{2,W}(\{s_i\}) v(Y) - \sum_i \mu_i^2 a_i^1 \pi_i U_{C^2}^{2,m} (1 + r),$$

with $\Phi_v^{2,W}(\{s_i\}) = \Phi_v^{2,m}(\{s_i\}) \sum_i \omega_y^i(\{s_i\}) \left[\frac{1}{\varphi_i^2} + \mu_i \gamma \right] \pi_i$. Besides, $\sum_i \mu_i^2 a_i^1 \pi_i$ no longer drops out of the Lagrangian and affects the social marginal utility of aggregate consumption at time 2. Using (1.28) and the expression derived before for $U^{2,m}$, we arrive at the optimal $\hat{\tau}_y$ in Proposition 1.4.

Shifting to the first period, note first that the first order condition for $(s_i)_{\forall i}$ in the fictitious representative consumer problem may be used to derive $U_S^{1,m} = \beta \rho \tilde{\Phi}_v^{2,m} v_Y(\tilde{Y}) \frac{\tilde{Y}}{S^{\gamma\rho+1}}$, which is the same we obtain under commitment except that \tilde{Y} and $\tilde{\varphi}_i^2$ replace Y and φ_i^2 . Also, a_i^1 is derived from the individual Euler equation (1.5) with zero marginal taxes; i.e., $c_i^1 = \tilde{c}_i^2$. For given private expectations on future policy under time-consistency, the first period allocation is decentralizable if

$$\begin{aligned} -\frac{U_S^{1,m}}{U_S^{2,m}} s_i - a_i^1 + a_i^0 + T^1 &= a_i^1 (1+r) + \tilde{T}^2 + y_i (1 - \tilde{\tau}_y) \\ \Leftrightarrow \\ a_i^1 &= \frac{a_i^0 + T^1 - \beta \gamma \rho \tilde{\Phi}_v^{2,m} \frac{v_Y(\tilde{Y})}{S^{\gamma\rho}} \omega_s^i C^1 - \tilde{T}^2 + \tilde{\omega}_y^i \tilde{Y} (1 - \tilde{\tau}_y)}{2+r}. \end{aligned}$$

An envelope theorem applied to the second period social programme with Lagrangian (1.27) defines

$$\begin{aligned} V_A^{2,soc} &= V_B^{2,soc} = \mathcal{W}_{C^2}^2 (1+r), \\ V_S^{2,soc} &= \frac{\partial \bar{W}^2}{\partial S} - \sum_i \mu_i^2 \frac{da_i^1}{dS} \pi_i U_{C^2}^{2,m} (1+r), \\ V_{C^1}^{2,soc} &= - \sum_i \mu_i^2 \frac{da_i^1}{dC^1} \pi_i U_{C^2}^{2,m} (1+r), \end{aligned}$$

with \bar{W}^2 standing for $W^2(\cdot)$ after incorporating $\{s_i\}_{i \in I}$ as functions of S . The full derivatives of $\mathcal{W}^1(\cdot)$ may thus be computed after replacing the individual allocations as dependent on aggregates and market weights and utilizing the last expressions. Finally, noting that $U_{C^1}^{1,m} / \beta \tilde{U}_{C^2}^{2,m} = 1+r$ as in the commitment case, that $\mathcal{W}_{C^1}^1 / \mathcal{W}_A^1 = 1$ in the social optimum and that $\mathcal{W}_A^1 = \beta \tilde{V}_A^{2,soc} = \beta (1+r) \tilde{W}_{C^2}^2$, we get $\tilde{\xi} = 1$ in formula (1.30). Thus, $(1 - \hat{x}) = (1 - \tilde{\tau}_y) \frac{\tilde{W}_Y^2 / \mathcal{W}_S^1}{\tilde{U}_Y^{2,m} / U_S^{1,m}}$, which renders the expression for \hat{x} in Proposition 1.4 upon substituting the relevant derivatives and noting that the implementability condition $c_i^1 = \tilde{c}_i^2$ implies $C^1 = \tilde{C}^2$. ■

Optimal Tax / Subsidy Rates in Detail:

Following Atkeson, Chari, Kehoe (1999), we may develop the expressions in (1.25), (1.26), (1.28) and (1.30) further by writing the social marginal rates of substitution as: for the

commitment case,

$$\begin{aligned}\frac{W_Y}{W_{C^2}} &= \frac{U_Y^{2,m} \left\{ \sum_{i \in I} \left[\frac{U_Y^i}{\beta U_Y^{2,m}} + \mu_i \left(\frac{dy_i}{dY} + H_i^Y \right) \right] \pi_i \right\}}{U_{C^2}^{2,m} \left\{ \sum_{i \in I} \left[\frac{U_{C^2}^i}{\beta U_{C^2}^{2,m}} + \mu_i \left(\frac{dc_i^2}{dC^2} + H_i^{C^2} \right) \right] \pi_i \right\}}, \\ \frac{W_Y}{W_S} &= \frac{\beta U_Y^{2,m} \left\{ \sum_{i \in I} \left[\frac{U_Y^i}{\beta U_Y^{2,m}} + \mu_i \left(\frac{dy_i}{dY} + H_i^Y \right) \right] \pi_i \right\}}{U_S^{1,m} \left\{ \sum_{i \in I} \left[\frac{U_S^i}{U_S^{1,m}} + \mu_i \left(\frac{ds_i}{dS} + H_i^S \right) \right] \pi_i \right\}},\end{aligned}$$

with

$$\begin{aligned}H_i^{C^2} &= \frac{\beta U_{C^2}^{2,m} c_i^2 + \frac{d(U_{C^1}^{1,m} c_i^1 + U_S^{1,m} s_i + \beta U_Y^{2,m} y_i)}{dC^2}}{\beta U_{C^2}^{2,m}}, \\ H_i^Y &= \frac{\beta U_{YY}^{2,m} y_i + \frac{d(U_{C^1}^{1,m} c_i^1 + \beta U_{C^2}^{2,m} c_i^2 + U_S^{1,m} s_i)}{dY}}{\beta U_Y^{2,m}}, \\ H_i^S &= \frac{U_{SS}^{1,m} s_i + \frac{d(U_{C^1}^{1,m} c_i^1 + \beta U_{C^2}^{2,m} c_i^2 + \beta U_Y^{2,m} y_i)}{dS}}{U_S^{1,m}}.\end{aligned}$$

If all agents are equal to the representative consumer, these expressions have an exact parallel with those derived by Atkeson, Chari and Kehoe (1999). Therein, the " H " terms are interpreted as general equilibrium elasticities, which is what planner cares about. In the case of no commitment¹¹,

$$\begin{aligned}\frac{\mathcal{W}_Y^2}{\mathcal{W}_{C^2}^2} &= \frac{U_Y^{2,m} \left\{ \sum_{i \in I} \left[\frac{U_Y^{2,i}}{\beta U_Y^{2,m}} + \mu_i^2 \left(\frac{dy_i}{dY} + \mathcal{H}_i^Y \right) \right] \pi_i \right\}}{U_{C^2}^{2,m} \left\{ \sum_{i \in I} \left[\frac{U_{C^2}^{2,i}}{\beta U_{C^2}^{2,m}} + \mu_i^2 \left(\frac{dc_i^2}{dC^2} + \mathcal{H}_i^{C^2} \right) \right] \pi_i \right\}}, \\ \frac{\widetilde{\mathcal{W}}_Y^2}{\mathcal{W}_S^1} &= \frac{\beta \widetilde{U}_Y^{2,m} \left\{ \sum_{i \in I} \left[\frac{\widetilde{U}_Y^{2,i}}{\beta \widetilde{U}_Y^{2,m}} + \mu_i^2 \left(\frac{d\widetilde{y}_i}{dY} + \widetilde{\mathcal{H}}_i^Y \right) \right] \pi_i \right\}}{U_S^{1,m} \left\{ \sum_{i \in I} \left[\frac{w'(c_i^1) \frac{dc_i^1}{dS}}{U_S^{1,m}} + \mu_i^1 \left(\frac{ds_i}{dS} + \mathcal{H}_i^S \right) \right] \pi_i + \frac{\beta \widetilde{V}_S^{2,soc}}{U_S^{1,m}} \right\}},\end{aligned}$$

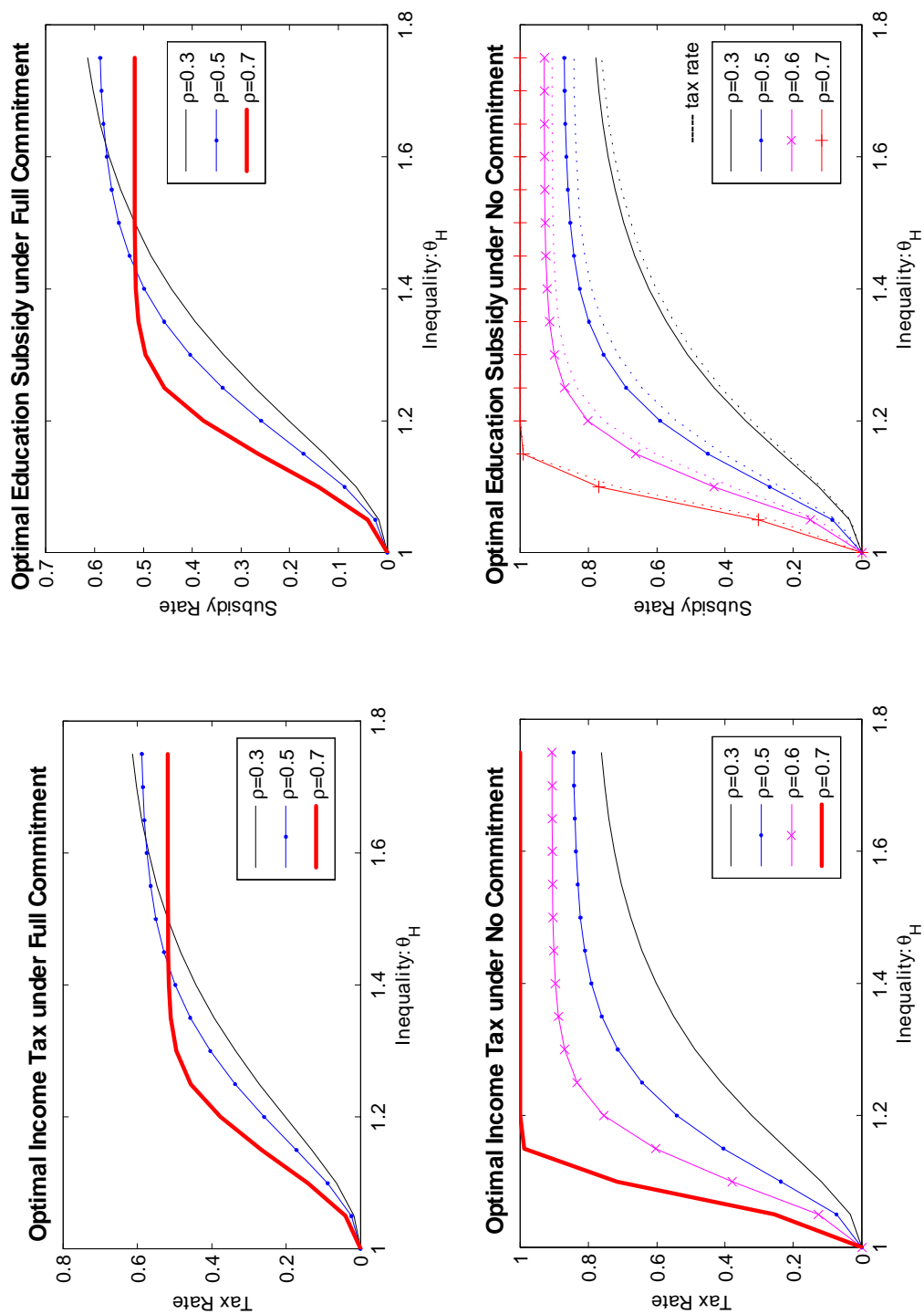
¹¹Noting that optimality with respect to T^1 implies $\sum_{i \in I} \mu_i^1 \pi_i = 0$, i.e., (1.29) does not depend on T^1 .

with

$$\begin{aligned}\mathcal{H}_i^{C^2} &= \frac{U_{C^2 C^2}^{2,m} c_i^2 + \frac{d(U_Y^{2,m} y_i)}{dC^2} - U_{C^2 C^2}^{2,m} a_i^1 (1+r)}{U_{C^2}^{2,m}}, \\ \mathcal{H}_i^Y &= \frac{U_{YY}^{2,m} y_i + \frac{d(U_{C^2}^{2,m} c_i^2)}{dY} - U_{C^2 Y}^{2,m} a_i^1 (1+r)}{U_Y^{2,m}}, \\ \mathcal{H}_i^S &= \frac{U_{SS}^{1,m} s_i + \frac{d(U_{C^1}^{1,m} (c_i^1 + a_i^1))}{dS} - U_{C^1 S}^{1,m} a_i^0}{U_S^{1,m}}.\end{aligned}$$

The social marginal rate of substitution between aggregate consumption and leisure in period 2 differs from the commitment to the no commitment case because the general equilibrium elasticities change. That is both due to a neglect of past incentives by the uncommitted government and the adverse/positive impact on second period implementability constraints through the $\sum_{i \in I} U_{C^2}^{2,m} a_i^1 (1+r) \pi_i$ term. The committed government disregards the latter because it faces an intertemporal implementability constraint and is free to set any intertemporal price.

In turn, the expression for $\frac{\widetilde{W}_Y^2}{W_S^1}$ is distinct from that of $\frac{W_Y}{W_S}$ to the extent that *i*) Y is set by the second period government not in the same way as the first period government would prefer, which in turn impacts the relative preference for higher schooling; *ii*) S - and through it also $\{a_i\}_{i \in I}$ - become state variables in the second period - S does not affect future policy rules but it does determine the scope for high second period (and, hence, lifetime) welfare. Finally, *iii*) policy makers cannot optimize over τ_a^0 . Hence it is relevant how far S affects implementability constraints through the $\sum_{i \in I} U_{C^1}^{1,m} a_i^0 (1+r) \pi_i$ term.



1.9 References

- Albanesi, S., Sleet, C., (2006). Dynamic Optimal Taxation with Private Information. *Review of Economic Studies* (73).
- Atkeson, A., Chari, V.V., Kehoe, P., (1999). "Taxing Capital Income: A Bad Idea," Federal Reserve Bank of Minneapolis Quarterly Review, vol. 23.
- Boadway, R., Marceau, N., Marchand, M., (1996). Investment in Education and the Time Inconsistency of Redistributive Tax Policy. *Economica* Vol.63.
- Bovenberg, L., Jacobs, B., (2005). Redistribution and Education Subsidies are Siamese Twins. *Journal of Public Economics* (89).
- Bovenberg, L., Jacobs, B., (2008). Optimal Taxation of Human Capital and the Earnings Function. CESifo Working Paper No. 2250.
- Bohacek, R., Kapicka, M., (2008). Optimal Human Capital Policies. *Journal of Monetary Economics*, vol. 55 (1).
- Carneiro, P., Heckman, J., (2003). Human Capital Policy. IZA Discussion Paper No. 821.
- Carneiro, P., Meghir, C., Parey, M., (2007). Maternal Education, Home Environments and the Development of Children and Adolescents. CEPR DP6505.
- Cunha, F., Heckman, J., Lochner, L., (2005). Interpreting the Evidence on Life Cycle Skill Formation. NBER wp 11331.
- Dur, R., Teulings, C. (2004). Are Education Subsidies an Efficient Redistributive Device?. Tinbergen Institute Discussion Paper 2003-024/3.
- Konrad, K. (2001). Privacy, Time-Consistent Optimal Labor Income Taxation and Education Policy. *Journal of Public Economics* 79(3).
- Kapicka, M., (2006). Optimal Income Taxation with Human Capital Accumulation and Limited Record Keeping. *Review of Economic Dynamics* (9).
- Mirrlees, J., (1971). An Exploration in the Theory of Optimal Income Taxation. *Review of Economic Studies* (38).
- OECD (2007). Education at a Glance. OECD Publishing.
- Pereira, J., (2008). Time-Consistent Redistribution Policy and Education Subsidies under Alternative Fiscal Constitutions., mimeo.
- Reis, C. (2005). Education and Optimal Taxation. mimeo, MIT.
- Werning, I., (2007). Optimal Fiscal Policy With Redistribution. *Quarterly Journal of Economics* (August 2007).

CHAPTER 2

TIME-CONSISTENT REDISTRIBUTION POLICY AND EDUCATION SUBSIDIES UNDER ALTERNATIVE FISCAL CONSTITUTIONS

2.1 Introduction

Classical theories of optimal fiscal policy often assume that the planner is restricted to the use of distortionary taxes to finance public expenditures. In this framework, the seminal contribution of Fischer (1980) has shown that second best allocations are not implementable if governments lack commitment, as it is always optimal to revise plans after investment decisions are made. If governments may resort to lump sum taxation, however, there is no time-inconsistency problem whatsoever and the first best might always be implemented. Mirrless (1971) builds up the case for distortionary taxation when governments care for equity but cannot observe all relevant individual characteristics. If agents differ in the ability to pay taxes with the same effort, such governments would like to use ability-contingent tax bills and fully redistribute *luck*. That is generally not incentive feasible if abilities are not observable; but by using distortionary taxation the government may explore differences in relative preferences for effort and provide partial insurance. Naturally, any information on the individual that is correlated with its ability will optimally be used by the policy maker. Like in the classical optimal taxation theory, the objective of the government is always to provide public goods - including redistribution - in the least distortionary way.

The central friction in the Mirrless framework is that the source of inequality is exogenous and unobservable. Arguably, though, labor productivity is not simply determined by *luck* but is (also) the product of voluntary investment in human capital carried out in the early stages of the life cycle. When governments have that into account, the optimal redistributive allocation is best implemented not only through income taxation but also with the use of education policy, so long as (at least part of) education investment is measurable by the fiscal authority. Bovenberg and Jacobs (2005), Reis (2005) and Maldonado (2007) have shown that to be the case in static settings with endogenous labor productivity. In general, distortionary subsidies to education should be used to offset indirect distortions caused by the taxation of labor supply.

There has been a recent revival of Mirrless's theory of optimal taxation as economists seek to extend its basic insights into dynamic problems. Important contributions in this literature include Golosov, Kocherlakota and Tsyvinski (2003), Kocherlakota (2005), Albanesi and Sleet (2005) and Werning (2007)¹. Bohacek and Kapicka (2008) consider again the case where skills are built up endogenously as the product of initial ability and time dedicated to human capital accumulation. Education policy plays a small role in their setup only because its costs are implicitly deductible in the tax bill already.

Importantly, as we frame the redistribution problem into a dynamic setting the time-inconsistency problem of second best policies reemerges. To see it, suppose agents types are permanent. After a first period has passed in which the government screens agents through the tax system, informational asymmetries are gone and the government may implement a first best redistribution policy. For the particular case of endogenous labor productivity, differences in early investment in education incentivated by an appropriate government subsidy schedule would allow perfect identification of individuals as they enter in the labor market. Income could, thus, be fully redistributed, which in this extreme case completely eliminates incentives for the most able individuals to study more a priori.

In practice, the disincentive will be as strong as tighter the relation between future income taxes and the worker educational status [presuming there is a correlation between investment and attainment] and that is bound to influence decisively the optimal design of education policy. With that in mind, in this paper we use a two-period, two-class economy with endogenous skills borrowed from Chapter 1 to compare optimal non-linear redistribution policy under two kinds of fiscal constitution: one in which the income tax schedule may depend on all observable individual variables, including past schooling; and another in which taxes may depend on the gross income of the individuals only, i.e., there is *limited record keeping*. Hence, we take in the latter a looser interpretation of the Mirrless (1971) informational frictions: even if governments could measure individual productivity, they are not allowed to condition income taxes on that information. Such is the property of the income taxes codes in virtually all real economies, which constitutes the main motivation for undergoing our proposed analysis.

We find that constraining the income tax function to depend only on earned income enhances social welfare when education cannot be subsidized at more than 100% of its cost. Intuitively, under full record second period governments always choose to fully redistribute labor income if they can condition income taxes on agents' productivities, for which past education is a perfect proxy in our model. Thus, returns on schooling are expropriated and

¹See also Golosov, Tsyvinsky and Werning (2007) for a comprehensive survey.

individuals are actually worse off for becoming more productive since they are *forced* later to work more. When taxes depend exclusively on gross income, incentives for labor supply are provided by allocating higher net income to the most productive individuals. Education has then a strictly positive marginal value for private agents. In equilibrium, aggregate resources increase sufficiently that, despite the absence of full redistribution, average utility is higher.

The case for limited record keeping is less clear if governments can effectively pay agents to study. But even in that case, the benefits from having perfect redistribution of resources during adulthood do not necessarily compensate the lifetime costs of necessary education distortions in the first period. That depends on how close we would be to perfect redistribution under limited record, which in turn relies on the parameters of the model. Anyway, we learn with our analysis that the implications for optimal education policy of one or the other fiscal constitutions are rather different. If only informational frictions bind and policy is chosen sequentially, there is a case of massive subsidization of non-compulsory education, even in the absence of any externality motive. By contrast, if the tax bill may depend solely on gross income, agents are optimally called to support part of their marginal education costs; in our stylized model more so for lower levels of education.

Time-inconsistency effects in non-linear redistribution policy with endogenous productivities have previously been discussed by Boadway et al. (1996). The authors consider an economy where governments may always observe final productivities, regardless of the commitment technology, but not education investment. Mandatory education is then shown to be welfare improving, a result that our solution for time-consistent policy under full record keeping generalizes. In a follow up paper, Konrad (2001) argues that, indeed, letting governments collect information on agents final productivity after schooling choices are made always hampers social welfare. Like in Boadway et al. (1996), the mechanism through which the government would acquire such information is not endogenous. Individuals are *ex ante* equal in his model, meaning that they choose equal amounts of schooling, if any. Our model is, thus, more general than these two previous analysis. We consider individuals who are differently able *a priori* and whose schooling choices are the source of full information in the second period. We also distinguish the inconsistency problem associated with information screening from that arising, *à la* Fischer (1980), with the intertemporal investment in human capital: skills are endogenous *ex-ante*, but inelastic *ex-post*. Finally, we allow for more general education policies in the intertemporal policy mix.

Our paper also relates to a recent body of literature on time-consistent redistribution. This literature is pioneered by Roberts (1984) who studies optimal income taxation with discrete types in infinite horizon, concluding that the only possible equilibrium is the pooling

of types by the fiscal authority. Berliant and Ledyard (2005) consider optimal nonlinear taxation in a two period economy with a continuum of types and permanent skills. Brett and Weymark (2008) also focuses on a two-period deterministic setting, but takes only two individuals and analyses the time-consistent redistribution problem when agents are allowed to save. A similar kind of problem is discussed by Bisin and Rampini (2006), who with a series of simple two-period, two-types examples show that granting private agents access to anonymous capital markets may improve social welfare. The argument for implementing limited record keeping in our paper is particularly related to their proposal.

By looking at a single life-cycle, we exclude all kinds of reputation mechanisms. In the context of the new dynamic public finance literature, Acemoglu et al. (2008) and Fahri and Werning (2008) have looked at the redistribution policies that such mechanisms can sustain in infinite horizon, also having into account political frictions. As expected, disincentives for truth-telling and investment are considerably lessened in that case, even if taxes may be history-dependent.

The paper is organized as follows. Section 2.2 sets up the economy, the individuals' problem and the equilibrium under *laissez faire*. In Section 2.3 we describe the first best allocation as a benchmark and in Section 2.4 the second best outcome. We argue in this section that limited record keeping is harmful when governments have commitment but works as a powerful commitment device otherwise, given the inconsistency problem inherent to the second best plan. The time-consistent policy mix is scrutinized in Section 2.5, emphasizing the contrast between the full and limited record solutions. Section 2.6 concludes the paper and Section 2.7 gathers the proofs.

2.2 The Economy

We consider a simple deterministic economy that lasts 2 periods: youth (y) and adulthood (o). There is a continuum of agents, divided into two types, $i = 1, 2$, which have equal weight in the population [the measure of the population is normalized to 2]. Types are defined by a different ability to profit from schooling, which we assume to be higher for type 1 than for type 2. Each agent of type i starts youth with cognitive ability θ_i ($\theta_1 > \theta_2$) and has then to decide how much to invest in education out of a net initial budget that we define bellow. The relative price of schooling is normalized to 1, so that s_i units of forgone consumption yield a future hourly wage equal to $\psi(\theta_i, s_i) = \psi_i$, with $\psi(\cdot)$ a C^2 function, strictly increasing and concave in both arguments and such that $\psi(\theta_i, 0) = \underline{\psi}$ and

$\lim_{s \rightarrow 0} \psi_s = +\infty$.² One may interpret the parameter θ as the joint product of innate ability, parents' early investment and nurture during childhood. We therefore assume that there is complementarity between schooling and cognitive ability: $\psi_{s\theta} > 0$.

During adulthood, agents earn gross labor income $Y_i = \psi_i n_i$, where n_i is the number of hours they choose to work out of a total time endowment \bar{n} . Both types have lifetime preferences given by

$$U(c_i^y, c_i^o, n_i) = u(c_i^y) + \beta [u(c_i^o) - v(n_i)], \quad (2.1)$$

with $0 < \beta < 1$, $u', v', v'' > 0$ and $u'' < 0$.

Individuals are endowed in the first period with a non-negative amount of resources $k_0/2$, which, without loss of generality, we assume to be equal across types. Importantly, they do not have access to capital markets. Although restrictive, this assumption notably simplifies the analysis and, for a reasonable parameterization of $\psi(\cdot)$, is equivalent to the usual premise that young agents cannot borrow against their future earnings in the private market. We comment further on the practical motivation for this restriction in footnotes 7 and 19.

The government in this economy simply designs a redistributive tax/subsidy system, having no net revenue requirements. Unlike young households, it may borrow/save in international capital markets at gross rate $1 + r$. Hence, it is partially able to alleviate the borrowing constraints faced in the private sector. Policy makers cannot generally to identify agents by their types, as θ_i , ψ_i and n_i are not directly observable, but know agents' preferences and the human capital technology and may observe total income, schooling choices and consumption levels. Fiscal policy is thus an intertemporal subsidy/tax schedule $\{S_i, T_i\}_{i=1,2}$ which may depend on *observable* variables only.³ That said, we will consider in this paper different scenarios corresponding to different fiscal constitutions. These amount to distinct set of rules on the particular form that $\{S_i, T_i\}_{i=1,2}$ may take.

For all cases, the individual budget constraints in the first and second periods are given by

$$c_i^y + s_i \leq k_0/2 + S_i, \quad (2.2a)$$

$$c_i^o \leq Y_i - T_i, \quad i = 1, 2. \quad (2.2b)$$

Agent i 's optimization problem is therefore that of maximizing (2.1) subject to these budget constraints and a time constraint $n_i \leq \bar{n}$, taking as given the redistribution policy⁴ and the

²The last condition is sufficient for an interior choice of s_i , $i=1,2$ in most equilibria considered in this paper, but it is not necessary.

³Taxes will not depend on consumption levels directly because those are completely determined by income, the tax bill and the education expenditure.

⁴Whilst in scenarios where the government may credibly commit to future policy T_i is exactly known

choices of all other agents. Instead, the government maximizes an utilitarian social welfare function subject to aggregate feasibility, information problems, commitment technology and the fiscal constitution.

It is shown in the appendix that the solution to the household's programme, when interior, is described by the first order conditions⁵

$$MRS_{s_i, c_i^y} : \frac{\beta v'(n_i) n_i \psi_s(\theta_i, s_i)}{\psi(\theta_i, s_i) u'(c_i^y)} = 1 - S_s(s_i) + \frac{\beta u'(c_i^o)}{u'(c_i^y)} T_s(s_i, Y_i), \quad (2.3)$$

$$-MRS_{Y_i, c_i^o} : \frac{v'(n_i)}{\psi(\theta_i, s_i) u'(c_i^o)} = 1 - T_Y(s_i, Y_i). \quad (2.4)$$

These are not affected by the timing of actions, as there is no uncertainty in the model and the households' choices are not subject to a time inconsistency problem. Equation (2.4) is standard; it simply equates the marginal rate of substitution of leisure for consumption in the second period to the net wage rate of effective labor supply [the gross wage rate of effective labor supply - $n_i \psi_i$ - is equal to 1]. (2.3), instead, states that the marginal rate of substitution of schooling for consumption in the first period, which for a given labor supply is decreasing in s_i , is higher if subsidies to education are low and/or if returns to education are marginally taxed ex-post. Notice that (2.3) is equivalent to

$$\frac{\beta u'(c_i^o)}{u'(c_i^y)} n_i \psi_s(\theta_i, s_i) [1 - T_Y(s_i, Y_i)] = 1 - S_s(s_i) + \frac{\beta u'(c_i^o)}{u'(c_i^y)} T_s(s_i, Y_i), \quad (2.5)$$

suggesting that *ceteris paribus* investment in education is discouraged by high income taxation.

We close this section with a description of a useful benchmark:

The laissez faire equilibrium

In *laissez faire* all taxes and marginal taxes are set to zero. In this case, both margins come undistorted, and so does the intertemporal margin $\beta u'(c_i^o)/u'(c_i^y)$. In particular, we have

$$\beta n_i \psi_s(\theta_i, s_i) = \frac{u'(k_0/2 - s_i)}{u'(n_i \psi(\theta_i, s_i))},$$

implying that the relation between cognitive ability and education is not necessarily monotonic, even in *laissez faire*. Applying the implicit function theorem to the individual problem, it is

at the beginning of lifetime, whenever policy is chosen sequentially agents must form expectations on the future tax schedule. It is an equilibrium requirement that such expectations are correct. Nonetheless, we are abusing notation by writing T_i irrespectively of the timing assumption of government choice. We do it as it comes at no interpretation cost.

⁵As is well know, in optimal non-linear taxation models with a finite number of types, the solution for the tax schedule is usually a non-differentiable (step) function. S_s, T_s and T_Y in equations (2.3) and (2.4) shall be interpreted as *implicit* marginal tax rates at the equilibrium and not as actual derivatives.

possible to show that for a sufficiently low elasticity of substitution between schooling and ability in the human capital technology and an intertemporal elasticity of substitution close to 1, we obtain $s_1 > s_2$ in the optimum [see also Boadway, Marceau and Machand (1996) for a environment with inelastic labor supply]. In effect, such seems to be the relevant case in real economies, but it is not a necessary prediction of our setup.

Non participation in capital markets implies that agents do not equally smooth consumption intertemporally.

2.3 Redistribution Policy with Complete Information

Suppose the government can recognize types, in the sense that it observes individual θ 's and may choose any $\{S_i, T_i\}_{i=1,2}$ measurable on θ_i . We follow the leading case in the literature and consider an utilitarian social welfare function

$$\begin{aligned} W^y(c_1^y, c_2^y, c_1^o, c_2^o, n_1, n_2) &= U(c_1^y, c_1^o, n_1) + U(c_2^y, c_2^o, n_2) \\ &= u(c_1^y) + u(c_2^y) + \beta W^o(c_1^o, c_2^o, n_1, n_2) \end{aligned} \quad (2.6)$$

with $U(\cdot)$ as defined in (2.1) and therefore $W^o(c_1^o, c_2^o, n_1, n_2)$ representing the social continuation value after the first period. The problem of the policy maker can be stated as one of choosing directly allocations $\{c_i^y, s_i, c_i^o, Y_i\}_{i=1,2}$ subject to a feasibility constraint instead of the tax system that implements them:

$$\begin{aligned} &\max_{\{c_i^y, s_i, c_i^o, Y_i\}_{i=1,2}} W^y(c_1^y, c_2^y, c_1^o, c_2^o, n_1, n_2) \\ &s.t. \\ &\sum_{i=1}^2 c_i^y + s_i + \frac{c_i^o}{1+r} \leq k_0 + \sum_{i=1}^2 \frac{Y_i}{1+r}, \end{aligned} \quad (2.7a)$$

$$Y_i = \psi(\theta_i, s_i) n_i, \quad n_i \leq \bar{n}, \quad i = 1, 2, \quad (2.7b)$$

$$k_0 \text{ and } \theta_1, \theta_2 \text{ are given.} \quad (2.7c)$$

We characterize its unique solution in Proposition 1.

Proposition 1: *At the first best allocation of the economy described in Section 2, (i) consumption is constant across types - $c_1^y = c_2^y$, $c_1^o = c_2^o$, (ii) there is perfect consumption smoothing - $u'(c_i^y)/u'(c_i^o) = \beta(1+r)$, $\forall i$ and (iii) schooling investment and labor supply are higher for the most able individual - $s_1 > s_2$ and $n_1 > n_2$. Therefore, type-1 agents are strictly worse-off than type-2 agents.*

Proof: See appendix.

Like in *laissez faire*, all marginal rates of substitution —and in this case also the intertemporal margin— are equal to 1. Thus, marginal taxes/subsidies must be null: $S_s = T_s = T_Y = 0$. Implementing the optimal allocation through a tax system is otherwise straightforward, since agents are exactly identifiable. The fiscal authority will just need to impose a *prohibitive* tax on choices other than the social optimum. Because there are no revenue requirements, we have that $S_1 + (1+r)^{-1}T_1 = -[S_2 + (1+r)^{-1}T_2] > 0$.

Noteworthy, the intertemporal composition of individual tax bills is indeterminate; only its present value matters. This is because in a first best equilibrium the absence of informational problems at the beginning of each period and the fact that the social planner does not need to resort to distortionary taxation put off time inconsistency problems. The solution to the problem is the same whether decisions are all taken at date zero or sequentially in the life cycle. It also does not depend on the fiscal authority keeping record of first period allocations, as the second period government is anyway fully informed of agents' types.

2.4 Optimal Nonlinear Redistribution Policy under Commitment

It follows from Proposition 1 that when the fiscal authority cannot observe θ the first best allocation is not incentive compatible. Agents *born* with high cognitive ability would, if they can, claim to be less talented and see their utility thereby increase. Therefore, we characterize in this section the mechanism design problem faced by the social planner under asymmetric information. Similar policy problems have been studied by Boverberg and Jacobs (2005, 2008), Reis (2005) and Bohacek and Kapicka (2008) for continuum of types, and Maldonado (2007) in a two-class economy. Since for now we keep the credibility assumption, a Revelation Principle allows us to focus on direct truth-telling mechanisms through which the government chooses report-contingent allocations which are incentive compatible. The planner solves:

$$\begin{aligned} & \max_{\{c_i^y, s_i, c_i^o, Y_i\}_{i=1,2}} W^y(c_1^y, c_2^y, c_1^o, c_2^o, n_1, n_2) \\ & s.t. \\ & u(c_i^y) + \beta[u(c_i^o) - v(n_i)] \geq u(c_j^y) + \beta \left[u(c_j^o) - v\left(\frac{Y_j}{\psi(\theta_i, s_j)}\right) \right], \quad i, j = 1, 2, \quad i \neq j \\ & \text{and conditions (2.7a), (2.7b), (2.7c).} \end{aligned} \tag{2.8}$$

Unlike in a two-class economy with an utilitarian social welfare function and separable utility, one cannot guarantee that only the incentive compatibility constraint of the highest type binds, unless we assume that $n_1 \geq n_2$ and $s_1 \geq s_2$ in the optimum [sufficient but

not necessary]. For illustrative purposes we will characterize these type of equilibria only throughout the various sections of the paper—the "normal" case, in the terminology of Stiglitz (1982) and a standard practice in the literature—, calling the attention of the reader whenever that option is consequential.

The solution to this pure second best policy problem is derived in the appendix and has similar properties to those found in the related literature:

Proposition 2: *At a second best allocation of the economy described in Section 2, (i) consumption is always higher for type 1 - $c_1^y > c_2^y$, $c_1^o > c_2^o$, (ii) there is perfect consumption smoothing - $u'(c_i^y)/u'(c_i^o) = \beta(1+r)$, $\forall i$, (iii) there is no distortion at the top: - $MRS_{Y_1, c_1^o} = MRS_{s_1, c_1^y} = 1$, (iv) labor supply of the low types is distorted downwards: $0 < -MRS_{Y_2, c_2^o} < 1$ and (v) net distortion on human capital investment depends on $\psi(\cdot)$: $0 < MRS_{s_2, c_2^y}$ is $<$, $=$ or $>$ than $-MRS_{Y_2, c_2^o}$ whenever ψ_s/ψ is increasing, constant or decreasing in θ , respectively.*

Apart from (i) and (iii), this is a general result, not confined to the "normal" case. The two conditions, as well as condition (iv), follow from the assumption that the most able individuals, and only those, benefit from an informational advantage in equilibrium. It is interesting to interpret (ii) and (v) at the light of such informational advantage.⁶ Whereas the policy maker can extract information on who is who due to an heterogeneous distaste for effective labor supply $Y_i = \psi_i n_i$ which induces agents to settle on a different marginal rate of substitution of consumption for leisure, intertemporal preferences for consumption are common due to separability. The rate at which they would optimally trade consumption intertemporally, had they the opportunity to do so, is exactly the same. Thus, the policy maker does not have any reason to give away this feature of the first best allocation. Clearly, investment in human capital is an indirect way of smoothing consumption across periods which individual agents can directly use, but in this case the government may be able to extract information by distorting the correspondent "intertemporal margin" $MRS_{s_2, c_2^y}/(-MRS_{Y_2, c_2^o})$, which it doesn't do in first best. Schooling expenditure may be understood as part of a *lifetime labor supply*; it reduces youth felicity through a drop in consumption but enhances productivity per hour or, equivalently, consumption per unit of foregone leisure in the future. If ψ_s/ψ is not constant in θ [elasticity of substitution in the human capital technology different from 1], the preference for more education holding all the other margins fixed is different, like the distaste for the observable Y_i is. Distorting this margin constitutes an additional mechanism of dealing with the information asymmetry between private individuals and the government. Namely, if s and θ are strong complements,

⁶For a deeper discussion of result (v) in a static environment see Bovenberg and Jacobs (2008).

s_2 is distorted downwards because type 1 likes schooling relatively more.⁷

As is well known, the decentralization of the second best equilibrium through a system of taxes and subsidies does not take a unique form. The allocations may be implemented with any fiscal plan that observes the implicit marginal taxes derived in the appendix and the feasibility constraints. In particular, the government may choose to set $T_s(i) = 0$, $i = 1, 2$ and $S_s(i) = 1 - MRS_{s_i, c_i^y}$ [or vice versa]. In that case, Proposition 2 implies $0 \leq S_s(i) < 1$, unless education is a very strong substitute for cognitive ability in the accumulation of human capital⁸. When ψ_s/ψ is constant in θ , marginal subsidies to education exactly match marginal income taxes. From (2.5), that corresponds to undistorted education investment, or a zero human capital wedge.

Second best equilibrium in the absence of education policy

Suppose governments may commit at date zero to a future income taxation policy but cannot set education subsidies/taxes —i.e., $\{s_i\}_{i=1,2}$ are not observable, meaning that in practice neither these nor consumption allocations are contractible by the fiscal authority. They would face a similar programme to (2.8), except that the incentive compatibility constraints are replaced by

$$u(c_i^y) + u(c_i^o) - v\left(\frac{Y_i}{\psi(\theta_i, s_i)}\right) \geq \max_{0 \leq \hat{s}_i \leq c_i^y + s_i} \left\{ u(c_i^y + s_i - \hat{s}_i) + u(c_j^o) - v\left(\frac{Y_j}{\psi(\theta_i, \hat{s}_i)}\right) \right\}, \quad i, j = 1, 2$$

involving an extra agency problem [notice that we do not require $i \neq j$] because agents may freely deviate from both planner's recommendations $(c_i^y, s_i)_{i=1,2}$. That is, investment in education is any individually affordable amount that follows equation (2.5) with $S_s(s_i) = T_s(s_i, Y_i) = 0$. In equilibria such that $\psi_1 \geq \psi_2$ it still holds that $T_Y(1) = 0$ and $T_Y(2) \in [0, 1]$ ⁹, and vice versa. Given our assumptions on the skill technology, this means that all agents still pick a strictly positive level of schooling, even if the implicit marginal cost is now higher for the less productive agent.

By construction, the income tax schedule initially chosen by the planner cannot depend on schooling in this case. We now turn to the discussion of equilibria with that same property but where governments are otherwise free to choose any education policy.

⁷It would be straightforward to allow for individual borrowing/saving in the second best. Apart from an additional non-arbitrage condition, the equilibrium properties would not be affected. As Brett and Weymark (2008) show in a related paper, the planner would not distort individual savings decisions precisely because agents have common intertemporal preferences for consumption.

⁸Empirical studies, e.g. after Carneiro and Heckman (2003), suggest a strong complementarity between cognitive ability built up during childhood and posterior investment on education.

⁹See proof of Proposition 1.

2.4.1 Nonlinear Policy with Limited Record Keeping: $T(s, Y) = T(Y)$

The inability of policy makers to condition tax plans on particular individual variables —be it innate characteristics such as θ in our model or choices like labor supply, schooling or consumption— is typically interpreted in the optimal non-linear taxation literature as the consequence of an information friction. In this section, we take the alternative view that the shape of a tax plan is surely conditioned by what the government can measure, but also by the fiscal constitution, i.e., a set of rules [tacit or explicit, in real economies] that dictate the variables upon which taxes can be made contingent. Naturally, this pulls us away from a pure Mirrleesian exercise. Nonetheless, policy makers are often *de facto* restricted on what they can tax/subsidize besides what they can effectively measure. It is appealing to ask what is the best tax schedule that observes such additional restrictions and to what extent social welfare is hampered [or not] by them.

In this paper, we are particularly interested in characterizing optimal redistribution policy when income taxation may not condition on past schooling choices¹⁰: $T(s, Y) = T(Y)$. Such form of limited record keeping has not been studied in the literature¹¹, despite its empirical relevance. As regards of the planner's problem stated in terms of quantities, it implies that equilibrium allocations must depend on reports of the respective period only. Therefore, in our economy the planner will have to elicit the true type of each agent twice.

In practice, the optimization programme is that of the second best policy (2.8) with two kinds of extra agency constraints. First, given first period truthful revelation, agents must have an incentive to keep on reporting their true type. I.e.,

$$u(c_i^o) - v(n_i) \geq u(c_j^o) - v\left(\frac{Y_j}{\psi(\theta_i, s_i)}\right), \quad i, j = 1, 2, \quad i \neq j. \quad (2.9)$$

Second, they must be deterred from mis-reporting in the first period only:

$$u(c_i^y) - \beta v(n_i) \geq u(c_j^y) - \beta v\left(\frac{Y_i}{\psi(\theta_i, s_j)}\right), \quad i, j = 1, 2, \quad i \neq j. \quad (2.10)$$

Importantly, even though agents are asked to reveal their types at different points in time, the tax authority commits at date zero to the announced fiscal plan. For parameterizations of the economy such that we fall in a "normal" case, we show in the appendix to Proposition 3 that this planner faces only two binding constraints: (2.9) for $i = 1$ and (2.10) for $i = 2$.

Proposition 3: *The limited record plan of a committed government implies, (i) consumption higher for type 1 during adulthood - $c_1^o > c_2^o$ - but lower during youth - $c_2^y > c_1^y$ (ii) imperfect*

¹⁰Kapicka (2006) and Gaube (2007) look at optimal non-linear taxation in economies where there is limited record of past incomes.

¹¹Certainly, limited record is an implicit assumption of all linear taxation models - see e.g. Pereira (2008). The extra degree of freedom we are allowing for is the shape of the tax function.

consumption smoothing - $u'(c_2^y)/u'(c_2^o) < \beta(1+r) < u'(c_1^y)/u'(c_1^o)$, (iii) no distortion of labor supply at the top: $MRS_{Y_1, c_1^o} = 1$, but (iv) $MRS_{s_1, c_1^y} < 1$, if $\frac{\partial(\psi_s/\psi)}{\partial\theta} \geq 0$ [indeterminate otherwise], (v) labor supply of the low types distorted downwards: $MRS_{Y_2, c_2^o} < 1$, but (vi) $MRS_{s_2, c_2^y} = 1$.

Proof: See appendix.

Equilibrium properties are significantly different from those in Proposition 2, hinting a lower achieved social welfare. Namely, new intertemporal distortions arise. Perfect consumption smoothing no longer holds, not because agents now have different intertemporal preferences over consumption but because informational advantages are not constant over time. In particular, in equilibria with $\psi_1 > \psi_2$ it is the group of highly able individuals that enjoys an informational advantage in the second period; but, for the same c^y , low ability individuals always benefit from higher levels of education - namely s_1 - if they are then free to report being of low type - work to generate Y_2 - in the second period. Naturally, low ability individuals would not prefer studying more during youth if that required having to produce as much as Y_1 in the future [as it does in the second best]. The necessary effort for that would not compensate mis-reporting during youth. Likewise, high ability agents would be less attracted by type 2s' second period allocation if they had to also pick (c_2^y, s_2) .

Hence, low types have an effective informational advantage in the first period vis-a-vis the planner, and $c_2^y > c_1^y$ is necessary to elicit types in the first period, whilst $c_1^o > c_2^o$ as usual. Providing redistribution entails poorer smoothing of consumption. Furthermore, education investment is now distorted upwards for type 1 [unless $\frac{\partial(\psi_s/\psi)}{\partial\theta}$ is really very small; see appendix] and downwards for type 2s, who see their income taxed at the margin in the last period. That holds even with an unitary elasticity of complementarity between θ and s . The perfect match between marginal subsidies and taxes identified in Bovenberg and Jacobs (2005) does not survive the limited record assumption because schooling and labor supply no longer affect incentives in a symmetric fashion.

The expressions provided in the appendix also suggest that type-2's labor wedge is widened —marginal taxes on income are higher—, although that depends on the equilibrium values for gross income of the two types. The new effect comes from the harmful impact that Y_2 has on the first period screening problem. Having a larger marginal tax indirectly reduces incentives to mimic type 1 during youth.

Evidence that the ex-ante *representative* household must be [weakly] worse off is simply the fact that the best limited record allocation was necessarily feasible for the second best planner - that of Proposition 2 - whilst the contrary is not true. That, of course, relies on the assumption that fiscal planners can commit up-front to the income tax schedule. When

governments lack a commitment technology one cannot draw the same conclusion, as we will make clear in Section 2.5.

We close this Section with a discussion of the commitment problem that will arise in our intertemporal redistribution setting.

2.4.2 Time Inconsistency and the Ratchet Effect

The second best equilibrium summarized in Proposition 2 suffers from a credibility problem. The reason is two-fold. First, there is a Fischer(1980)-like time inconsistency issue associated with the optimal distortions over human capital investment. Once youth has passed and schooling choices are made, differences in skill levels are fixed and gross income may be redistributed more effectively by disregarding the effects on past incentives to study —i.e., part of *lifetime labor supply* is seen as inelastic. Second, at the beginning of adulthood the asymmetric information problem is already resolved because youth allocations are type-dependent. A benevolent policy maker re-optimizing at this date would better use all the available information and eliminate distortions, giving rise to a ratchet effect whereby the government may extract more income taxes from the most able individuals if they reveal their types up-front.

When the government is free to condition the tax schedule on whatever available information, these two problems collapse indeed into a single time-inconsistency problem; that of the ratchet effect. Recent contributions by Berliant and Leyard (2005), Bisin and Rampini (2006), Brett and Weymark (2008) and Krause (2008) deal with the time inconsistency of nonlinear taxation in two-period economies under such approach¹². To the best of our knowledge, there is not yet a study of time-consistent redistribution policy in economies where the ratchet effect is avoided by the fiscal constitution but that still face a commitment problem in the nonlinear policy design. Bisin and Rampini (2006) consider the potential benefits of granting agents access anonymous markets and eliminating the ratchet effect. However, the economies under study are such that anonymous markets also prevent *any* redistribution in the second period [there will simply be no observable basis for taxation].

To see how the ratchet effect operates in our economy consider the outcome of Section 2.3. In particular, consider the optimality conditions for labor supply and consumption in the second period. When second period governments know the identity of agents, net incomes are equalized and the most skilled individuals enjoy relatively less leisure than in the second best. That is, all their information rents are lost. Hence, they face a strong disincentive

¹²Freixas, Guesnerie and Tirole (1985) and Dillén and Lundholm (1996) analyse the ratchet effect in the context of optimal linear taxation in two period settings.

to reveal their type a priori, be it through a higher schooling investment or any distinctive choice that tags them as the most skilled. In fact, when second period governments are fully informed, the marginal value of education is negative from an individual perspective. Because there's a continuum of individuals, future consumption is taken as given, whilst labor supply is strictly increasing in productivity.

Time inconsistency in the absence of education policy: the case for limited record keeping

The distortion over schooling is especially severe in the absence of education policy or if the fiscal authority for some reason cannot propose marginal subsidies to education above 100%, which would not bind under commitment. Consider the incentives faced by a single individual when second period governments observe schooling choices. If all other individuals choose levels of schooling such that $s_1 \neq s_2$ [i.e., there's a first best in the second period], she is always better off dropping investment in education to zero, consuming more in the first period and consuming the same and working less in the following period [because the impact of her deviating on total resources would be negligible]. Thus, there is no possible equilibrium with $s_1 \neq s_2$. When $0 < s_1 = s_2$ future governments remain uninformed of types and therefore implement a traditional second best allocation for the one last period. Nonetheless, individuals of type 2 at least would also prefer not to study at all and being identified with the skill level $\psi(\theta, 0) = \underline{\psi}$. That would actually lead to higher consumption not only in the first but also in the second period¹³ and, again, to less working hours. Naturally, because all type 2 individuals behave symmetrically, we would be back to a $s_1 \neq s_2$ situation. Thus, as long as individuals face a weakly positive marginal price of schooling and there is full record keeping with no government commitment investment in education is zero for all types¹⁴. Social welfare would be lower than in *laissez faire* as there would be no redistribution at all in either period and both agents would be constrained in their choice of schooling¹⁵. Being lower than in *laissez faire*, it is also necessarily lower than the second best equilibrium because the former is a feasible allocation for the second best planner.

Limited record keeping is undoubtedly beneficial in this case. Even in the absence of any education policy, if governments cannot condition income taxes on previous schooling choices,

¹³See proof of Proposition 1 in the appendix.

¹⁴Konrad (2001) discusses a similar result in an economy where final differences in productivity are random and not type specific.

¹⁵Boadway et al (1996) characterize a similar situation with no education policy [we refer to the time-consistency formulation of Section IV in their paper]. Their economy has a finite number of agents of each type, but again if this number is sufficiently large, incentives to educate are so low that no one does it. The mechanism is slightly different from ours because the authors assume that productivities are always observable in the second period. There is no ratchet effect, then, but the effect of time inconsistency is still the same.

second period redistribution policy is set as in a one-shot two-class economy [*a la* Stiglitz (1982)] where ψ_i is fixed but is private information. The ratchet effect is no longer into play as governments now *commit* not to explore acquired information on individual types. Therefore $0 \leq T_Y(i) < 1$, inducing agents to invest a strictly positive share of their resources in education as was the case in the full commitment equilibrium without education subsidies. Yet, the endogeneity of schooling choices —i.e., their cost in terms of youth consumption— is still disregarded by the policy maker at the start of adulthood. Consequentially, the income tax schedule is suboptimal from the first period point of view; a time-inconsistency problem subsisting despite the abolition of the ratchet effect.

Theorem 1: *In the absence of education policy, or if implicit marginal subsidies to education are constrained to a 100% ceiling, full record keeping and lack of commitment induce an equilibrium with no education investment and social welfare below that of the laissez faire. In this scenario, limited record keeping (i) improves lifetime social welfare, but (ii) is not sufficient to restore the equilibrium with perfect commitment. The lack of credibility is, thus, still welfare detrimental.*

Proof: See appendix.

As regards (i), for consumption in both periods and leisure as normal goods, the increase in social welfare happens both *ex ante* and *ex post* education decisions. (ii) asserts the distinction between time-inconsistency in the choice of the tax schedule and the pure ratchet effect that we discussed above. Welfare is lower than the equilibrium described in Proposition 3, to which the only difference is the commitment assumption. By transitivity, it is also necessarily worse than the pure second best allocation.

When schooling is observable, it is straightforward for any government to rule out a zero schooling situation by simply imposing a sufficiently heavy tax on choices below the minimum optimal level of s . The problem then becomes the asymmetric information on who is the cleverest *a priori*. In the next section we characterize time-consistent redistribution policy when the government is free to set any incentive scheme for education, comparing solutions and social welfare under the full and limited record regimes.

2.5 Optimal Time-Consistent Nonlinear Redistribution Policy

In this last section we study the properties and implementation of socially optimal allocations chosen sequentially in time. Our main objective is to contrast the solution under the realistic assumption that income tax schedules are set contingent on gross earnings only—in our model, only labor earnings although what we intend to stress is the separation of $T(\cdot)$ from the education level—and under no *ad hoc* restriction other than the initial asymmetric

information over θ_i , $i = 1, 2$.

We start by the latter. As has been recognized in the literature, when governments cannot credibly commit not to exploit in the future collected information on individual types, application of the standard Revelation Principle is precluded. Disincentives to reveal one's type a priori may be so large that the first period policy maker may prefer to resort to non-revealing [pooling] mechanisms or to mixed strategies.¹⁶ Comparison across the best of the different options is not clear cut as it depends on the parameters of the model. In all, the government will balance the benefits from full insurance in the second period under separation with those accruing from lower distortions over human capital investment under pooling.

We will focus on symmetric equilibria, in which all agents of the same type choose the same allocations when faced with the same incentives. These are either separating or pooling equilibria. The study of semi-pooling equilibria is beyond the scope of this paper.

2.5.1 Separation in Youth with Full Insurance in Adulthood

When the first period fiscal plan leads the two types to choose different $(c_i^y, s_i)_{i=1,2}$ allocations, second period governments are perfectly informed of individual ability to generate income. The redistribution problem is thus a first best:

$$\begin{aligned}
 V^{o,sep}(s_1, s_2, A) &= \max_{\{c_i^o, Y_i\}_{i=1,2}} W^o(c_1^o, c_2^o, n_1, n_2) \\
 s.t. & \\
 c_1^o + c_2^o &\leq A(1+r) + Y_1 + Y_2, \\
 Y_i &= \psi(\theta_i, s_i) n_i, \quad n_i \leq \bar{n}, \quad i = 1, 2, \\
 \psi(\theta_1, s_1), \psi(\theta_2, s_2), A &\text{ are given.}
 \end{aligned} \tag{2.11}$$

Since $W^o(\cdot)$ is strictly concave and the constraint set is convex, the problem features a unique solution. This is an optimal allocation which is a function of the effective skill levels —themselves functions of $(\theta_i, s_i)_{i=1,2}$ — and inherited public savings/debt A . Therefore, indirect second period welfare may be written as $V^{o,sep}(s_1, s_2, A)$, which is continuous and differentiable in its arguments by an implicit function theorem. Note that (s_1, s_2) in $V^{o,sep}(s_1, s_2, A)$ stand for the human capital investment that all individuals of each type carry. Correspondingly, we may write the individual value function as $v^o(s_i)$, denoting the continuation value after investing s_i in education and taking as given the choices of all other

¹⁶See Bestzer and Strausz (2001, 2003).

agents and expectations on future tax rules. We summarize some properties of the solution to this last period redistribution problem in Lemma 1.

Lemma 1: *For each vector of state variables $(\psi(\theta_1, s_1), \psi(\theta_2, s_2), A)$, the unique second period equilibrium with separation of types satisfies: (i) $c_1^o = c_2^o = c^o$, $n_1 > (\text{or } =) n_2$ iff $\psi(\theta_1, s_1) > (\text{or } =) \psi(\theta_2, s_2)$ (ii) $-MRS_{Y_i, c_i^o} = 1$, $i=1,2$, (iii) $V_A^{o,sep} = u'(c^o)(1+r)$, (iv) $V_{s_i}^{o,sep} = u'(c^o) n_i \psi_s(\theta_i, s_i) > 0$ and (v) $v_{s_i}^{o,sep} = -v'(n_i) \frac{\partial n_i}{\partial \psi_i} \psi_s(\theta_i, s_i)$.*

Conditions (i) and (ii) follow from Proposition 1 and the irrelevance of a commitment technology in the equilibrium described therein. The policy maker is able to provide full redistribution with no efficiency loss. In practice, this means that taxes are non-distortionary and type specific. In other words, fiscal policy in this period is a tax rule $\{T(s_i, Y_i)\}_{i=1,2}$ such that $T_Y(i) = 0, \forall i$.

Parts (iii) and (iv) are obtained with an envelope theorem. The impact on social continuation value of an extra unit of schooling for all individuals of each type $[V_{s_i}^{o,sep}]$ is naturally positive. It is higher for the types with higher labor supply and/or marginal return on schooling in terms of labor productivity. Contrarily, for the individual agent studying has negative value when it implies working more in the future. (v) was already derived in the proof of Theorem 1 and identifies schooling as a "bad" for individuals when productivity is observable in the second period.

At the start of the life-cycle both the agents and the government are aware of the incentives of the future government. Lemma 1 is common knowledge. What is more, in Proposition 1 we asserted the willingness of the youth policy maker to induce agents to invest a positive share of their resources in education. Indeed, that is always feasible as we argued at the end of Section 2.4. Nevertheless, the first best allocation involving $c_1^y = c_2^y$ and $s_1 > s_2$ is not incentive compatible. Type 1 individuals have to be provided with incentives to reveal their types in the first period, given that they will be certainly worse off in the second period [in particular, worse off than by claiming to be of type 2]. That amounts to the incentive compatibility constraint

$$u(c_1^y) + \beta \left[u(c^o) - v \left(\frac{Y_1}{\psi(\theta_1, s_1)} \right) \right] \geq u(c_2^y) + \beta \left[u(c^o) - v \left(\frac{Y_2}{\psi(\theta_1, s_2)} \right) \right]. \quad (2.12)$$

Two important comments are in order. First, (2.12) is the relevant incentive constraint for truth telling insofar as agents disregard the effect of their mimicking on future aggregate resources and distribution of types. This is why c^o , Y_1 and Y_2 , as chosen by the future government, are simply taken as given in its expression. Second, in equilibrium c^o , Y_1 and Y_2 come out of the well behaved maximization problem (2.11) as continuous functions of the endogenous state variables (s_1, s_2, A) . When choosing these values in the present

[and note that we refer now to s_1 and s_2 as the "schooling policy" and not the particular investment on schooling by an agent], it is not indifferent for the policy maker the extend to which agents condition their expectations of c^o , Y_1 and Y_2 on (s_1, s_2, A) . Put differently, the intra-period timing of actions is consequential. If the government announces a tax schedule—or, equivalently, the triple (s_1, s_2, A) to be implemented—at date zero and commits to it within all the youth period, expectations of future allocation values depend on what has been announced. That would have to be considered in (2.12). But if the government has no "intra-period commitment", individual [rational] expectations over c^o , Y_1 and Y_2 depend only on initial state variables [distribution of cognitive abilities and initial resources] and not on current policy.¹⁷ Then, (2.12) shall be included in the first period social programme exactly as we have written it above. Alternatively, we may interpret that individuals are *future tax schedule* takers in the first period instead of *future tax rule* takers in this economy. Their forecasts on the credible future policy is then seen by the government as independent of first period policy.

Due to the convexity of $v(\cdot)$ and properties of the skill technology, in all equilibria where it is still optimal to induce $s_1 > s_2$ (2.12) is the only binding incentive constraint. In that case, the social problem at date zero is

$$\begin{aligned}
 & \max_{A, \{c_i^y, s_i\}_{i=1,2}} u(c_1^y) + u(c_2^y) + \beta V^{o,sep}(s_1, s_2, A) \\
 & s.t. \\
 & c_1^y + c_2^y + s_1 + s_2 + A \leq k_0, \\
 & \text{conditions (2.12) and (2.7c),} \\
 & V^{o,sep}(s_1, s_2, A) \text{ as defined in (2.11).}
 \end{aligned} \tag{2.13}$$

The government chooses the future endogenous state variables having into account how they affect future indirect social welfare, as expressed in Lemma 1. Optimality conditions differ from what we obtained in Section 3 as long as the respective variable relaxes or tightens the first period agency problem. That is particularly the case with (c_2^y, s_2) , which size relative to (c_1^y, s_1) determines how attractive it is for type 1 agents to misrepresent their type.

Proposition 4: *At a separating time-consistent equilibrium with $s_1 > s_2$ ¹⁸ : (i) $c_1^y > c_2^y$, (ii) $u'(c_1^y) < \beta(1+r)u'(c^o) < u'(c_2^y)$, (iii) $MRS_{s_1, c_1^y} = 1$ and (iv) $MRS_{s_2, c_2^y} < 1$ if*

¹⁷Cohen and Michel (1988) were the first to point out this difference within the literature of time-consistent fiscal policy. See Ortigueira (2006) for a detailed study of intra-period timing assumption in the Markov-perfect taxation of capital.

¹⁸The agency problem may lead to an optimal allocation with $s_2 > s_1$. Even in this case, $c_1^y > c_2^y$ is still necessary in all separating equilibria. The final solution is such that the main insights on how this equilibrium compares to both the pooling solution of Section 5.2 and limited record in Section 5.3. are not particularly affected. Thus, we leave out the full characterization of this case.

$\frac{\partial(\psi_s/\psi)}{\partial\theta} \leq 0$. Parts (i) and (ii) of Lemma 1 apply. Lifetime utility is higher for type 1 agents: $U(c_1^y, c_1^o, n_1) > U(c_2^y, c_2^o, n_2)$.

Proof: See appendix.

By (2.12) and Lemma 1, highly able individuals will invest more in human capital only if their consumption during youth is also higher. Since the incentive problem is irrelevant in the second period, perfect consumption smoothing is impossible if types are to be separated in the first period. As in the second best equilibrium, the policy maker screens agents through the differences in their relative preferences for schooling. There is no distortion at the top, but individuals of type 1 have a stronger taste for schooling at type-2's allocation than have agents of type 2. When the skill technology is weakly separable — $\frac{\partial(\psi_s/\psi)}{\partial\theta} = 0$ —, it is then optimal to introduce a positive marginal subsidy to education at the equilibrium s_2 level. As in the second best, the case for a marginal subsidy is stronger if education and ability are weak complements. If, on the contrary, the elasticity of complementarity in the skill technology is higher than 1 — i.e., $\frac{\partial(\psi_s/\psi)}{\partial\theta} > 0$ — the optimal subsidy is reduced and it may even turn optimal to tax education for agent 2.

Implementation

We now discuss the implementation of the time-consistent separating equilibrium through a system of government prices. From part (v) of Lemma 1 and the discussion in Section 2.4.2, we know that all agents resort to zero education if they have to support part of the investment costs. One may reinterpret this in terms of the implicit fiscal distortion. Redistribution with full efficiency is obtained during adulthood because the tax schedule is history dependent. Namely, the income tax schedule does not vary with Y_i but with s_i : $T_i = T(\theta_i, s_i)$, s.t. $T_1 + T_2 = -A(1+r)$ and $T_1 > (or =) T_2$ iff $\psi_1 > (or =) \psi_2$. It is straightforward to show that $T(\theta_i, s_i)$ is increasing in the individual s_i . The reason is twofold: all returns from schooling are expropriated because consumption is equalized ex post and the future policy maker will *force* the agent to work more with the rise of her individual productivity. There is no counteracting income effect because the weight of her income in total resources is negligible. Differentiating the individual budget constraint (2.2b) with respect to s_i at the adulthood social equilibrium we obtain

$$T_s(\theta_i, s_i) = \left. \frac{\partial n_i}{\partial s_i} \right|_{\bar{c}^o} \psi(\theta_i, s_i) + n_i \psi_s(\theta_i, s_i),$$

where \bar{c}^o is the expected future per capita consumption, which the individual takes as given. It is shown in the appendix to Theorem 1 that this derivative, at the eyes of the individual agent, depends only on \bar{c}^o and ψ_i . Denote it $n_s(\bar{c}^o, \psi_i)$. Replacing $T_s(s_i, Y_i)$ by the expression

above and $T_Y(s_i, Y_i)$ by 0 in (2.5) we get that an interior choice of schooling s_i may only be implemented if marginal subsidies to education observe

$$S_s(s_i) \geq 1 + n_s(\bar{c}^o, \psi_i) \psi(\theta_i, s_i) \frac{\beta u'(c^o)}{u'(c_i^y)}.$$

Therefore, at the optimal allocation the government has to be subsidizing agents at the margin in more than 100% to achieve its preferred "schooling policy" (s_1, s_2) . Agents have to be compensated for the whole cost of additional studying plus the discounted value of the extra taxes they end up paying as adults for having studied more in the present. This may be interpreted either as an actual payment of the education activity or as a potential loss in transfers if education investments fall below the social optimum.

From Proposition 4 we know that there is no marginal subsidy at the top. There is a positive one at the bottom as long as the elasticity of complementarity between education and ability is low enough [although it still can be higher than 1]. In that case, the lifetime social optimal allocation under separation, full record keeping and lack of commitment may be implemented with the subsidy schedule:

$$S(s) = \begin{cases} \Omega^{low}(s), & \text{if } s \leq s_2 \\ \Omega^{high}(s), & \text{otherwise} \end{cases},$$

with $\Omega^{low}(s)$ and $\Omega^{high}(s)$ such that:

$$\begin{aligned} \Omega_s^{low}(s) &= 1 + n_s(\bar{c}^o, \psi_2) \psi(\theta_2, s) \frac{\beta u'(c^o)}{u'(c_2^y + s_2 - s)}, & \text{if } s \leq s_2 \text{ and } \Omega_s^{low}(s) = 0 \text{ otherwise,} \\ \Omega_s^{high}(s) &= 1 + n_s(\bar{c}^o, \psi_1) \psi(\theta_1, s) \frac{\beta u'(c^o)}{u'(c_1^y + s_1 - s)}, & \text{if } s \leq s_1 \text{ and } \Omega_s^{high}(s) = 0 \text{ otherwise.} \end{aligned}$$

c_1^y, c_2^y, s_1, s_2 are the solutions to (2.13). With this schedule the government again *forces* agents to choose the optimal allocation. Both the present and future costs of investing in education are just exactly covered so that agents are indifferent between investing one more unit in education or not. Noteworthy, the incentive has to be provided in the first period, as in the second period governments would simply renege on any promised transfer conditional on past education.

2.5.2 Pooling in Youth with Second-Best in Adulthood

When governments can keep record of previous choices, the only circumstance in which the second period government remains uninformed of individual types arises if in the first

period consumption and education choices are common across all agents; e.g., if the government simply mandates a minimum level of education (and we know already that no one has an incentive to study more than that if not paid to do so). We now study the properties of the best equilibrium with such property. For that, as we have done in the previous subsection, we start by analyzing the second period social programme with inherited endogenous variables (\bar{s}, A) :

$$\begin{aligned}
 V^{o,pool}(\bar{s}, A) &= \max_{\{c_i^o, Y_i\}_{i=1,2}} W^o(c_1^o, c_2^o, n_1, n_2) \\
 s.t. & \\
 c_1^o + c_2^o &\leq A(1+r) + Y_1 + Y_2, \\
 u(c_1^o) - v\left(\frac{Y_1}{\psi(\theta_1, s)}\right) &\geq u(c_2^o) - v\left(\frac{Y_2}{\psi(\theta_1, s)}\right), \\
 Y_i &= \psi(\theta_i, \bar{s}) n_i, \quad n_i \leq \bar{n}, \quad i = 1, 2, \\
 \psi(\theta_1, \bar{s}), \psi(\theta_2, \bar{s}), A &\text{ are given.}
 \end{aligned} \tag{2.14}$$

Since $\psi(\theta_1, \bar{s}) > \psi(\theta_2, \bar{s})$, $\forall \bar{s}$, only one incentive compatibility constraint binds; that of type 1. Furthermore, knowing \bar{s} and θ_1, θ_2 , the government may infer the distribution of skills at the beginning of the working life. The solution to (2.14) has thus the canonical properties of the one-period two class economy:

Lemma 2: *For each pair of endogenous state variables (\bar{s}, A) , at a second period equilibrium with previous pooling of types satisfies: (i) $c_1^o > c_2^o$, (ii) there is no distortion at the top: $-MRS_{Y_1, c_1^o} = 1$, but n_2 is distorted downwards: $-MRS_{Y_2, c_2^o} < 1$, (iii) $u'(c_1^o)(1+r) < V_A^{o,pool} < u'(c_2^o)(1+r)$ and (iv) $V_{\bar{s}}^{o,pool} > \sum_{i=1}^2 v'(n_i) n_i \frac{\psi_s(\theta_i, \bar{s})}{\psi(\theta_i, \bar{s})}$.*

Parts (i) and (ii) are standard, whilst parts (iii) and (iv) follow from an envelope argument applied to programme (2.14). Notice that for the same total consumption during the second period, the value of savings is now lower because there is no perfect equalization of net resources. On the other hand, a marginal increase in the investment in education—in this case, of all agents necessarily—has the additional benefit of relaxing the agency problem during adulthood.

In the first period the government chooses (\bar{c}^y, \bar{s}, A) having into account how the endogenous state variables affect the continuation value. Since allocations are common to all

agents, there is no extra agency problem and the problem faced by the government is simply

$$\begin{aligned}
 & \max_{(\bar{c}^y, \bar{s}, A)} 2u(\bar{c}^y) + V^{o,pool}(\bar{s}, A) \\
 & s.t. \\
 & 2\bar{c}^y + 2\bar{s} + A \leq k_0, \\
 & \text{condition (2.7c),} \\
 & V^{o,pool}(\bar{s}, A) \text{ as defined in (2.14).}
 \end{aligned} \tag{2.15}$$

Due to the infinite productivity of the first units of education and the fact that increasing \bar{s} relaxes second period incentive constraint, the government always opts for a positive \bar{s} in this equilibrium. Following our discussion from Section 2.4.2, also in this case subsidies to education have to be such that agents are implicitly paid to study at the margin, above the pure schooling cost. Such implementation is more direct for the pooling equilibrium as the policy maker would just have to impose a high enough tax on schooling choices different from \bar{s} . Note that this is, indeed, a non-revealing tax schedule.

The properties of the equilibrium are otherwise indeterminate. It is possible to show that the marginal rate of substitution between human capital investment and youth consumption is distorted for at least one of the agents. If $n_1 > n_2$ and ψ_s/ψ is increasing in θ , one may also conclude that the implicit MRS_{s,c^y} is higher for the high types and necessarily below 1 for the low types; but nothing further.

As in the separating equilibrium, lifetime utility is higher for the most able, as they consume the same in the first period and enjoy information rents in the second.

2.5.3 Optimal Time-Consistent Nonlinear Redistribution Policy under Limited Record Keeping

In this last subsection we determine optimal schooling and income taxation policy when the latter cannot be history dependent by a *constitutional* constraint. We argue this to be a relevant study case insofar as it is virtually the practice in all real economies: tax codes do not rely on information the government could use about individuals' schooling levels, even if there is public financial support for non-compulsory education.

As above, the redistribution problem is solved backwards. By construction, the ratchet effect is discarded and we may exclude non-revealing mechanisms.¹⁹ Despite the government

¹⁹For this end the restriction of no individual borrowing/saving is crucial. Even if income taxes cannot depend on schooling, when savings are taxable - that is, observable - it is possible to infer agent's types. Moreover, as long as savings differ across agents, the government will fully redistribute them in our model creating an incentive to borrow infinitely.

not being authorized to condition adults' allocation on previous schooling, we continue to assume that in equilibrium policy makers know the correct distribution of skills as of the start of the second period. The problem it solves is a simple one-period Mirrleesian programme with asymmetric information on the skills ψ_1 and ψ_2 . This problem is stated in the proof of Theorem 1. With no surprise, when $\psi_1 > \psi_2$ the equilibrium shares properties (i) to (iii) of Lemma 2.²⁰ There is no distortion at the top, but type 2 agents will face a positive marginal tax rate. Highly skilled agents consume more and, as usual, derive higher utility in equilibrium. Also, part (iv) of Lemma 2 is valid for s_1 , that is, higher education for the most able relaxes the second period incentive compatibility constraint. This is not true of s_2 .

In what follows we focus on equilibria comprising $n_1 \geq n_2$, which is sufficient but not necessary to ensure $s_1 > s_2$ is the preference of the youth policy maker²¹. Suppose first that during youth the government was fully informed of individual cognitive abilities. It would then choose an allocation such that both individuals would consume the same. Is this allocation incentive compatible? To answer this question, consider first that an agent considering misrepresenting her type neglects the impact of her action on the future distribution of skills perceived by the policy maker. The future allocation —i.e., pairs $(c_i, Y_i)_{i=1,2}$ from which to self-select— is thus expected to be the same. Only there is now the possibility of choosing different type reports in different periods. That said, it is clear that type 1 agents prefer their allocation to that of type 2 agents. With the same c^y , they would just become less productive and derive strictly less lifetime utility under any adult skill-type report. However, low type agents would profit from claiming to have cognitive ability θ_1 in the first period and then picking the (c_2^o, Y_2) pair as adults²². Denoting $V^{o,lr}(s_1, s_2, A)$ the indirect utility in the second period under limited record keeping, we may define the lifetime redistribution

²⁰It is possible to show that first period policy makers will never set s_1, s_2 so that $\psi_1 < \psi_2$. The proof consists in: (i) identifying all the relevant incentive compatibility constraints [for both periods] as we do in this section for the "normal" case and (ii) showing that there is an alternative allocation inducing $\psi_1 \geq \psi_2$ which liberates resources for consumption for the same social desutility of effort and which weakly relaxes the agency problem.

²¹When $n_2 > n_1$ first period wedges have a less clear pattern than that conveyed in Proposition 5 below. Nonetheless, the comparison across fiscal constitutions is analogous.

²²See proof of Proposition 3.

problem as that of solving

$$\begin{aligned}
& \max_{A, \{c_i^y, s_i\}_{i=1,2}} u(c_1^y) + u(c_2^y) + \beta V^{o,lr}(s_1, s_2, A) \\
& s.t. \\
& c_1^y + c_2^y + s_1 + s_2 + A \leq k_0, \\
& u(c_2^y) + \beta \left[u(c_2^o) - v\left(\frac{Y_2}{\psi(\theta_2, s_2)}\right) \right] \geq u(c_1^y) + \beta \left[u(c_2^o) - v\left(\frac{Y_2}{\psi(\theta_2, s_1)}\right) \right], \\
& \text{conditions (2.7c)}.
\end{aligned} \tag{2.16}$$

When type 2's incentive compatibility constraint binds type 1 agents still prefer their allocation, for which (2.16) comprises only one agency condition. This implies that $c_2^y > c_1^y$ in equilibria with $s_1 > s_2$, which contrasts with the time-consistent equilibrium with full record keeping and separation of types. Still, lifetime utility is higher for the highly able individuals than for those with lower cognitive ability.

In Proposition 5 we fully characterize the solution to (2.16).

Proposition 5: *At a time-consistent equilibrium with limited record such that $n_1 \geq n_2$: (i) $c_2^y > c_1^y$ whilst $c_2^o < c_1^o$, (ii) $U(c_1^y, c_1^o, n_1) > U(c_2^y, c_2^o, n_2)$, (iii) $-MRS_{Y_1, c_1^o} = 1$ whilst $MRS_{s_1, c_1^o} < 1$, if $\frac{\partial(\psi_s/\psi)}{\partial\theta} \geq 0$, indeterminate otherwise (iv) $-MRS_{Y_2, c_2^o} < 1$ whilst $MRS_{s_2, c_2^o} = 1$.*

Proposition 3 portrayed an equilibrium with exactly the same properties. This comes with no surprise, as the pattern of binding incentive compatibility constraints is the same. Similarly to the commitment solution, we also have here that for a non-empty interval of values below 1 for the elasticity of complementarity between θ and s in $\psi(\cdot)$, a positive marginal subsidy to s_1 is still optimal. Likewise, perfect consumption smoothing is impossible as it is optimal for agent 2 to dissociate reports in youth and adulthood.

The expressions for the second period wedges provided in the appendix portray the effect of lack of commitment under limited record. This is a Fischer (1980)-like distortion and translates into the second period government setting a more redistributive tax plan given the inherited state variables. That is, $-MRS_{Y_2, c_2^o}$ is *ceteris paribus* set closer to 1, the first best value. That likely implies a lower Y_2 and a tighter first period incentive compatibility constraint, leading to an optimally lower MRS_{s_1, c_1^o} [higher marginal subsidy].

Ideally, one would like to reach an unambiguous conclusion as to the (un)desirability of limited record when governments choose sequentially but education policy may take any form; as we did in Theorem 1 for education policies such that $0 \leq S_s(s_i) \leq 1$. However, even in our very stylized set up that is not possible. We would always need to resort to numerical comparison. The reason is that whilst the full record dully curbs incentives to study, when

the government is free to *pay* agents more than their investment cost, a minimum level of schooling may be implemented without distortion. Thus, only the agency problem is left and the loss from separating types is considerably lower. It may continue to be the case that the gains from full insurance in the second period outweigh the costs from distortions during youth. If that happens, limited record harms social welfare. Yet, take any parameterization of the model for which pooling is better than separation under full record. Then, limited record is certainly welfare improving, as a pooled solution is always feasible in the social programme (2.16).

Theorem 2: *If education policy is unrestricted when governments lack commitment, the limited record keeping is not necessarily a welfare improving fiscal constitution. There is, however, an open set of parameters for which the utilitarian social welfare is still higher under limited record - $T_i = T(Y_i)$ - than under full record - $T_i = T(s_i, Y_i)$. For all parameterizations, non-revealing mechanisms are necessarily suboptimal if limited record may be credibly implemented.*

Notice that the higher the complementarity between ability and education in the skill function, *ceteris paribus*, the more likely it is that the economy is better off restricting income taxes to depend on income only. Not only low type agents gain less from mimicking in the first period but also the second period allocation is closer to a first best. In other words, the same schooling levels (s_1, s_2) entails less distortionary costs of redistribution. As a consequence, the opportunity cost from not having a second period first best is lower.

Implementation

Like we did at the end of Section 5.1., we close our analysis with a proposal for the decentralization of the time-consistent equilibrium under limited record.

By assumption, income taxes must be irresponsive to the education levels. Furthermore, the tax schedule has to create the labor wedges enunciated in Proposition 5. As it is well known, the following tax system implements the social optimum (c_1^o, Y_1, c_2^o, Y_2) :

$$\tilde{T}(Y) = \begin{cases} T_{low}, & \text{if } Y \leq Y_2 \\ T_{high}, & \text{otherwise} \end{cases}.$$

T_{low} and T_{high} are set exactly so that the individual budget constraints are met with equality at the socially optimal allocation. No individual will have an incentive to choose anything other than the allocation intended for her by the policy maker. Like in the perfect redistribution equilibrium characterized in Lemma 1, it induces type 1 agents to pay higher taxes than type 2 agents, but consumption levels are not equalized.

Consider that $\frac{\partial(\psi_s/\psi)}{\partial\theta} \geq 0$. Then, optimal first period wedges are: for agents of type 1, $(1 - MRS_{s_1, c_1^o}) \in (0, 1)$; for agents of type 2, $1 - MRS_{s_2, c_2^o} = 0$. Since the marginal value of schooling is always strictly positive from an individual point of view [consider expression (2.5) with any non-negative $S_s(s_i)$ and $T_s(s_i, Y_i) = 0$], the policy maker does not need to *impose* each single unit of investment in human capital on the agent, as it did in Section 2.5.1. The subsidy schedule may actually take a similar form to that of the income tax one:

$$\tilde{S}(s) = \begin{cases} S_{low}, & \text{if } s < s_1 \\ S_{high}, & \text{otherwise} \end{cases}.$$

S_{high} is not necessarily higher than S_{low} [given the future tax schedule, high types do prefer their recommended allocation] but is certainly such that the most talented spend a bigger share of net resources on education. In fact, it is such that these individuals spend a higher share of their own initial resources on education in equilibrium.

Note that $\tilde{S}(s)$ is chosen such that the social optimum (c_1^y, s_1, c_2^y, s_2) is implemented and induces (c_1^o, Y_1, c_2^o, Y_2) in the future given the structure of $\tilde{T}(Y)$. By adjusting the subsidy schedule and therefore agents' optimal education investment, the first period government is actually indirectly choosing $\tilde{T}(Y)$, in both its cut off level and size of the transfers.

2.6 Conclusion

In this paper we considered a two-class economy where individual productivity is endogenous and redistribution policy is chosen sequentially. We showed that a fiscal constitution precluding the dependence of income taxes on past education investment may improve utilitarian social welfare. In particular, when governments are constrained to subsidize education at no more than 100% of its cost, the welfare gain is undoubtedly positive.

The limited record keeping institution would always be detrimental if governments could comply to their promised tax schedules after a big time span, as it limits the possible shape of fiscal instruments that can be used for redistribution. However, if governments reoptimize at the beginning of each period it works as a partial commitment device, completely eliminating the ratchet effect. In other words, it credibly stops future governments from exploiting information collected through the adverse selection of young agents.

We use our setup to identify the two potential sources of time-inconsistency in the optimal redistribution policy. Despite getting rid of the ratchet effect through the limited record institution, the benevolent policy maker still cannot restore the allocation it would choose *ex ante* under commitment. Future governments will ignore the impact of income taxation on past incentives to invest in human capital, perceiving a lower efficiency cost for the same

equity goals than they would *a priori*. That is an exact parallel to the time inconsistency problem identified by Fischer (1980) for homogeneous agents economies.

Throughout our analysis we assumed a perfect correlation between education investment and effective productivity enhancement for a given cognitive ability. Arguably, the relationship between the two is noisy, especially if what governments observe is only part of total schooling investment. However, as long as there is some positive association our main insights still go through. Measurable education expenditures will still be used as an efficient tagging device if governments are allowed to use it. Thus, individual agents are discouraged to raise them and that is harmful if these type of expenditures are strongly complementary to non-observable educational effort.

2.7 Appendix

Individual problem of an agent of type i

A general specification of the fiscal policy involves $S_i = S(s_i)$ and $T_i = T(s_i, Y_i)$. An individual agent of type i solves, then,

$$\begin{aligned} & \max_{\{c_i^y, s_i, c_i^o, Y_i\}} \left\{ u(c_i^y) + \beta \left[u(c_i^o) - v \left(\frac{Y_i}{\psi(\theta_i, s_i)} \right) \right] \right\} \\ & s.t. \\ & c_i^y + s_i \leq k_0/2 + S(s_i), \quad c_i^o \leq Y_i - T(s_i, Y_i), \\ & n_i = Y_i/\psi(\theta_i, s_i) \leq \bar{n}, \quad c_i^y, s_i, c_i^o, Y_i \geq 0, \\ & k_0 \text{ and } \theta_i \text{ are given.} \end{aligned}$$

Denoting $\lambda_{y,i}$ and $\lambda_{o,i}$ the multipliers for the budget constraints and assuming the solution is interior, the necessary conditions for an optimum are

$$\begin{aligned} & u'(c_i^y) - \lambda_{y,i} = 0 \\ & -\lambda_{y,i} (1 - S_s(s_i)) + \beta \left[v'(n_i) n_i \frac{\psi_s(\theta_i, s_i)}{\psi(\theta_i, s_i)} - \lambda_{o,i} T_s(s_i, Y_i) \right] = 0 \\ & u'(c_i^o) - \lambda_{o,i} = 0 \\ & -\beta v'(n_i) \frac{1}{\psi(\theta_i, s_i)} + \beta \lambda_{o,i} (1 - T_Y(s_i, Y_i)) = 0, \end{aligned}$$

which after rearranging terms yield expressions (2.3)-(2.4). ■

Proof of Proposition 1:

$W^y(c_1^y, c_2^y, c_1^o, c_2^o, n_1, n_2)$ is a strictly concave function and all the binding constraints of the first best programme are linear. Therefore, there is a unique solution to the problem. (i)

and (ii) follow directly from first order conditions and the separability of the utility function. To derive (iii) notice that the optimality condition for labor supply is $v'(n_i) = \lambda^* \psi(\theta_i, s_i)$ where λ^* is the multiplier of the feasibility constraint. Thus, by the convexity of $v(\cdot)$, $\psi_1 > (\text{or } =) \psi_2 \Rightarrow n_1 > (\text{or } =) n_2$.

We prove first that $n_1 > n_2$. Suppose not: $n_1 \leq n_2$. Then, it must be the case that $s_2 > s_1$ was chosen, otherwise it would be impossible to have $\psi(\theta_1, s_1) \leq \psi(\theta_2, s_2)$. Knowing how the optimal choice over labor supply hinges on the relationship between productivities, the government could set a s'_2 and s'_1 , such that $\psi(\theta_1, s'_1) = \psi(\theta_2, s_2)$ and $\psi(\theta_2, s'_2) = \psi(\theta_1, s_1)$. The continuation value would be the same and in the first period there would be resources relieved because $\psi_\theta > 0$ and $\psi_{s\theta} > 0, \forall s$. These resources can then be used to increase consumption and thus social welfare.

That said, it is left to show that $n_1 > n_2 \Rightarrow s_1 > s_2$. To check it, note that the first order condition with respect to schooling for each type i is $n_i \psi_s(\theta_i, s_i) = 1 + r$ - a non-arbitrage condition. Since $\psi_{ss} < 0$ and $\psi_{s\theta} > 0$, $n_1 > n_2 \Rightarrow s_1 > s_2$. ■

Proof of Proposition 2:

The Lagrangian for programme (2.8) is

$$\begin{aligned} \mathcal{L} = & u(c_1^y) + u(c_2^y) + \beta \left[u(c_1^o) - v\left(\frac{Y_1}{\psi(\theta_1, s_1)}\right) + u(c_2^o) - v\left(\frac{Y_2}{\psi(\theta_2, s_2)}\right) \right] + \\ & + \mu \left\{ u(c_1^y) + \beta \left[u(c_1^o) - v\left(\frac{Y_1}{\psi(\theta_1, s_1)}\right) \right] - u(c_2^y) + \beta \left[u(c_2^o) - v\left(\frac{Y_2}{\psi(\theta_1, s_2)}\right) \right] \right\} - \\ & - \lambda \left(c_1^y + c_2^y + s_1 + s_2 + \frac{c_1^o + c_2^o}{1+r} - k_0 - \frac{Y_1 + Y_2}{1+r} \right) \end{aligned}$$

yielding first order conditions for interior equilibria:

$$\begin{aligned} (c_1^y, c_2^y) : & u'(c_1^y)(1 + \mu) = u'(c_2^y)(1 - \mu) = \lambda \\ (c_1^o, c_2^o) : & \beta u'(c_1^o)(1 + \mu) = \beta u'(c_2^o)(1 - \mu) = \lambda / (1 + r) \\ (s_1) : & \beta v' \left(\frac{Y_1}{\psi(\theta_1, s_1)} \right) \frac{Y_1}{\psi(\theta_1, s_1)} \frac{\psi_s(\theta_1, s_1)}{\psi(\theta_1, s_1)} (1 + \mu) = \lambda \\ (s_2) : & \beta v' \left(\frac{Y_2}{\psi(\theta_2, s_2)} \right) \frac{Y_2}{\psi(\theta_2, s_2)} \frac{\psi_s(\theta_2, s_2)}{\psi(\theta_2, s_2)} \left(1 - \mu \frac{v' \left(\frac{Y_2}{\psi(\theta_1, s_2)} \right) \frac{1}{\psi(\theta_1, s_2)} \frac{\psi_s(\theta_1, s_2)}{\psi(\theta_1, s_2)}}{v' \left(\frac{Y_2}{\psi(\theta_2, s_2)} \right) \frac{1}{\psi(\theta_2, s_2)} \frac{\psi_s(\theta_2, s_2)}{\psi(\theta_2, s_2)}} \right) = \lambda \\ (n_1) : & \beta v' \left(\frac{Y_1}{\psi(\theta_1, s_1)} \right) \frac{1}{\psi(\theta_1, s_1)} (1 + \mu) = \lambda / (1 + r) \\ (n_2) : & \beta v' \left(\frac{Y_2}{\psi(\theta_2, s_2)} \right) \frac{1}{\psi(\theta_2, s_2)} \left(1 - \mu \frac{v' \left(\frac{Y_2}{\psi(\theta_1, s_2)} \right) \frac{1}{\psi(\theta_1, s_2)}}{v' \left(\frac{Y_2}{\psi(\theta_2, s_2)} \right) \frac{1}{\psi(\theta_2, s_2)}} \right) = \lambda / (1 + r) \\ \text{equat. (2.7a), } & \mu, \lambda > 0 \end{aligned}$$

$0 < \mu < 1$ for interior consumption allocations. Therefore, optimality conditions for consumption determine (i) and (ii). (iii) and (iv) follow from the conditions for (c_1^o, n_1, s_1)

and (c_2^o, n_2) , respectively, and (v) is dictated by the comparison of FOCs for labor supply and schooling for type 2.

Notice that $v' \left(\frac{Y_2}{\psi(\theta_2, s_2)} \right) \frac{1}{\psi(\theta_2, s_2)} > v' \left(\frac{Y_1}{\psi(\theta_1, s_1)} \right) \frac{1}{\psi(\theta_1, s_1)}$ in equilibrium. These quantities are equal in the first best, meaning that for the same skill levels $\psi(\theta_1, s_1)$ and $\psi(\theta_2, s_2)$, n_1/n_2 is now lower. ■

Proof of Proposition 3:

When types are to be elicited at the beginning of each period, three potential incentive compatibility constraints arise for each individual: mimicking in both periods, mimicking in the first period only and mimicking in the second period only [equat. (2.9)]. It follows that in "normal" redistributive equilibria with $\psi_1 \geq \psi_2$ and $Y_1 > Y_2$, (2.9) binds for $i = 1$ but not for $i = 2$ by the convexity of $v(\cdot)$. The former also implies that: (i) individuals of type 1 prefer mimicking in both periods than in the first period only and (ii) individuals of type 2 prefer mimicking in the first period only than in the two periods. Finally, for $s_1 \geq s_2$ only the incentive compatibility constraint prevailing in (ii) matters:

$$u(c_2^y) + \beta \left[u(c_2^o) - v \left(\frac{Y_2}{\psi(\theta_2, s_2)} \right) \right] \geq u(c_1^y) + \beta \left[u(c_2^o) - v \left(\frac{Y_2}{\psi(\theta_2, s_1)} \right) \right]. \quad (2.17)$$

When this condition exactly binds, we have necessarily that highly able individuals are not tempted to lie about their type in both periods. The reverse is not true; (2.17) always binds.

As such, the full programme of the government under a "normal" equilibrium has the associated Lagrangian

$$\begin{aligned} \bar{\mathcal{L}} = & u(c_1^y) + u(c_2^y) + \beta \left[u(c_1^o) - v \left(\frac{Y_1}{\psi(\theta_1, s_1)} \right) + u(c_2^o) - v \left(\frac{Y_2}{\psi(\theta_2, s_2)} \right) \right] + \\ & + \bar{\mu} \left\{ u(c_2^y) - \beta v \left(\frac{Y_2}{\psi(\theta_2, s_2)} \right) - \left[u(c_1^y) - \beta v \left(\frac{Y_2}{\psi(\theta_2, s_1)} \right) \right] \right\} - \\ & + \bar{\pi} \left\{ u(c_1^o) - v \left(\frac{Y_1}{\psi(\theta_1, s_1)} \right) - \left[u(c_2^o) - v \left(\frac{Y_2}{\psi(\theta_1, s_1)} \right) \right] \right\} \\ & - \bar{\lambda} \left(c_1^y + c_2^y + s_1 + s_2 + \frac{c_1^o + c_2^o}{1+r} - k_0 - \frac{Y_1 + Y_2}{1+r} \right), \end{aligned}$$

for which first order conditions at interior equilibria are

$$\begin{aligned}
(c_1^y, c_2^y) : u'(c_1^y)(1 - \bar{\mu}) &= u'(c_2^y)(1 + \bar{\mu}) = \bar{\lambda} \\
(c_1^o, c_2^o) : \beta u'(c_1^o)(1 + \bar{\pi}) &= \beta u'(c_2^o)(1 - \bar{\pi}) = \bar{\lambda}/(1 + r) \\
(s_1) : \beta v' \left(\frac{Y_1}{\psi(\theta_1, s_1)} \right) \frac{Y_1}{\psi(\theta_1, s_1)} \frac{\psi_s(\theta_1, s_1)}{\psi(\theta_1, s_1)} &\left[1 - \bar{\mu} \frac{v' \left(\frac{Y_2}{\psi(\theta_2, s_1)} \right) \frac{Y_2}{\psi(\theta_2, s_1)} \frac{\psi_s(\theta_2, s_1)}{\psi(\theta_2, s_1)}}{v' \left(\frac{Y_1}{\psi(\theta_1, s_1)} \right) \frac{Y_1}{\psi(\theta_1, s_1)} \frac{\psi_s(\theta_1, s_1)}{\psi(\theta_1, s_1)}} + \bar{\pi} \left(1 - \frac{v' \left(\frac{Y_2}{\psi(\theta_1, s_1)} \right) Y_2}{v' \left(\frac{Y_1}{\psi(\theta_1, s_1)} \right) Y_1} \right) \right] = \bar{\lambda} \\
(s_2) : \beta v' \left(\frac{Y_2}{\psi(\theta_2, s_2)} \right) \frac{Y_2}{\psi(\theta_2, s_2)} \frac{\psi_s(\theta_2, s_2)}{\psi(\theta_2, s_2)} &(1 + \bar{\mu}) = \bar{\lambda} \\
(Y_1) : \beta v' \left(\frac{Y_1}{\psi(\theta_1, s_1)} \right) \frac{1}{\psi(\theta_1, s_1)} (1 + \bar{\pi}) &= \bar{\lambda}/(1 + r) \\
(Y_2) : \beta v' \left(\frac{Y_2}{\psi(\theta_2, s_2)} \right) \frac{1}{\psi(\theta_2, s_2)} &\left[1 + \bar{\mu} \left(1 - \frac{v' \left(\frac{Y_2}{\psi(\theta_2, s_1)} \right) \frac{1}{\psi(\theta_2, s_1)}}{v' \left(\frac{Y_2}{\psi(\theta_2, s_2)} \right) \frac{1}{\psi(\theta_2, s_2)}} \right) - \bar{\pi} \frac{v' \left(\frac{Y_2}{\psi(\theta_1, s_1)} \right) \frac{1}{\psi(\theta_1, s_1)}}{v' \left(\frac{Y_2}{\psi(\theta_2, s_2)} \right) \frac{1}{\psi(\theta_2, s_2)}} \right] = \bar{\lambda}/(1 + r)
\end{aligned}$$

equat. (2.7a), $\bar{\mu}, \bar{\lambda}, \bar{\pi} > 0$

$0 < \bar{\mu}, \bar{\pi} < 1$ for interior consumption allocations, implying $c_1^y < c_2^y$ and $c_1^o > c_2^o$. Furthermore,

$$\frac{\beta(1+r)u'(c_1^o)}{u'(c_1^y)} = \frac{1-\bar{\mu}}{1+\bar{\pi}} \quad \text{and} \quad \frac{\beta(1+r)u'(c_2^o)}{u'(c_2^y)} = \frac{1+\bar{\mu}}{1-\bar{\pi}} \quad (2.18)$$

$$\frac{\beta v'(n_1) n_1 \psi_s(\theta_1, s_1)}{\psi(\theta_1, s_1) u'(c_1^y)} = \frac{1-\bar{\mu}}{1 - \bar{\mu} \frac{v' \left(\frac{Y_2}{\psi(\theta_2, s_1)} \right) \frac{Y_2}{\psi(\theta_2, s_1)} \frac{\psi_s(\theta_2, s_1)}{\psi(\theta_2, s_1)}}{v' \left(\frac{Y_1}{\psi(\theta_1, s_1)} \right) \frac{Y_1}{\psi(\theta_1, s_1)} \frac{\psi_s(\theta_1, s_1)}{\psi(\theta_1, s_1)}} + \bar{\pi} \left(1 - \frac{v' \left(\frac{Y_2}{\psi(\theta_1, s_1)} \right) Y_2}{v' \left(\frac{Y_1}{\psi(\theta_1, s_1)} \right) Y_1} \right)} \quad (2.19)$$

$$\frac{\beta v'(n_2) n_2 \psi_s(\theta_2, s_2)}{\psi(\theta_2, s_2) u'(c_2^y)} = 1 \quad (2.20)$$

$$\frac{\beta v'(n_2)}{\psi(\theta_2, s_2) u'(c_2^o)} = \frac{1-\bar{\pi}}{1 - \bar{\pi} \frac{v' \left(\frac{Y_2}{\psi(\theta_1, s_1)} \right) \frac{1}{\psi(\theta_1, s_1)}}{v' \left(\frac{Y_2}{\psi(\theta_2, s_2)} \right) \frac{1}{\psi(\theta_2, s_2)}} + \bar{\mu} \left(1 - \frac{v' \left(\frac{Y_2}{\psi(\theta_2, s_1)} \right) \frac{1}{\psi(\theta_2, s_1)}}{v' \left(\frac{Y_2}{\psi(\theta_2, s_2)} \right) \frac{1}{\psi(\theta_2, s_2)}} \right)} \quad (2.21)$$

$$\frac{\beta v'(n_1)}{\psi(\theta_1, s_1) u'(c_1^o)} = 1, \quad (2.22)$$

upholding Lemma 3, as for $s_1 > s_2$ and as long as $\frac{\partial(\psi_s/\psi)}{\partial\theta}$ is not too negative MRS_{s_1, c_1^y} and $-MRS_{Y_2, c_2^o}$ are lower than 1. ■

Proof of Theorem 1:

The zero education result under full record and $S_s(s_i) \leq 1$ follows from the discussion in the text. The optimality conditions in the second period under perfect information [e.g., when in the first periods agents chose $s_1 \neq s_2$] are such that $u'(c_i^o)$ is constant - $c_i^o = \bar{c}^o$, $\forall i$ - and $v'(n_i) = u'(c_i^o)\psi_i$ [see proof of Proposition 1]. What is more, an increase in income of a single individual has a negligible effect on total resources and consequentially on average consumption. Therefore, the ex ante marginal value of schooling from an individual perspective is

$$\left[u'(c_i^o) \frac{\partial c_i^o}{\partial \psi_i} - v'(n_i) \frac{\partial n_i}{\partial \psi_i} \right] \frac{\partial \psi_i}{\partial s_i} = -v'(n_i) \frac{\partial n_i}{\partial \psi_i} \frac{\partial \psi_i}{\partial s_i} < 0.$$

The social first order condition for individual labor supply n_i completely defines it as a function of a variable that by the beginning of adulthood is exogenous - ψ_i - for a given level of aggregate resources. We may, then, use an inverse function theorem to obtain: $\left. \frac{\partial n_i}{\partial \psi_i} \right|_{c_i^o = \bar{c}^o} = u'(\bar{c}^o)/v''(u'(\bar{c}^o)\psi_i) > 0$, sustaining the negative sign of the expression above. Since the marginal cost is weakly positive - $u'(c_i^o)(1 - S_s(s_i))$ - the best response of each individual agent to an expectation that others' choices are such that second period policy makers are fully informed is the corner solution $s_i = 0$. If $0 < s_i = s_2$, the second period government is uninformed of types [knows only the distribution of $\psi's$]. Take an individual agent of type 2 who chooses to instead set s to zero. Individuals with $s = 0$ are perfectly identifiable and therefore cannot mimic or be mimicked by other types with other levels of observable schooling. The F.O.C.s for consumption in an economy with the 3 types $\psi_1 > \psi_2 > \underline{\psi}$, in which only the incentive compatibility constraint of the highest type binds are such that $u'(c_1) < u'(c_{\underline{\psi}}) < u'(c_2)$, implying $c_2 < c_{\underline{\psi}}$. Moreover, we would have $v'\left(\frac{Y_{\underline{\psi}}}{\underline{\psi}}\right) \frac{1}{\underline{\psi}} < v'\left(\frac{Y_2}{\psi_2}\right) \frac{1}{\psi_2}$. Since $\underline{\psi} < \psi_2$, n_2 is necessarily higher than $n_{\underline{\psi}}$. Type 2 individuals would always prefer to change their education investment from s_1 to zero, meaning that we would be back to a $s_1 \neq s_2$ situation, which as shown above is not an equilibrium.

When there is limited record keeping but the policy maker cannot commit to future policy, income taxes are chosen so as to implement

$$\begin{aligned} \{c_1^o, c_2^o, n_1, n_2\}^* &= \arg \max W^o(c_1^o, c_2^o, n_1, n_2) \\ \text{s.t. equat. (2.9)} \\ c_1^o + c_2^o &\leq \psi(\theta_1, s_1)n_1 + \psi(\theta_2, s_2)n_2 + A(1+r) \end{aligned} \quad (2.23)$$

with $W^o(c_1^o, c_2^o, n_1, n_2)$ given as in (2.6) and $\psi(\theta_1, s_1)$, $\psi(\theta_2, s_2)$ not observable but which distribution is known to the policy maker. First order conditions for consumption and labor supply read [in equilibria with $\psi(\theta_1, s_1) \geq \psi(\theta_2, s_2)$, (2.9) binds only for $i = 1$]:

$$\begin{aligned} (c_1^o, c_2^o) : u'(c_1^o)(1 + \tilde{\mu}) &= u'(c_2^o)(1 - \tilde{\mu}) = \tilde{\lambda}, \quad \tilde{\mu} > 0 \\ (n_1) : \frac{v'\left(\frac{Y_1}{\psi(\theta_1, s_1)}\right)}{u'(c_1^o)\psi(\theta_1, s_1)} &= 1 = 1 - T_Y(Y_1) \\ (n_2) : \frac{v'\left(\frac{Y_2}{\psi(\theta_2, s_2)}\right)}{u'(c_2^o)\psi(\theta_2, s_2)} &= \frac{1 - \tilde{\mu}}{1 - \tilde{\mu} \frac{v'\left(\frac{Y_2}{\psi(\theta_1, s_1)}\right) \frac{1}{\psi(\theta_1, s_1)}}{v'\left(\frac{Y_2}{\psi(\theta_2, s_2)}\right) \frac{1}{\psi(\theta_2, s_2)}}} = 1 - T_Y(Y_2) < 1. \end{aligned} \quad (2.24)$$

We characterize the lifetime equilibrium with limited record keeping in the absence of education policy and show it renders higher welfare than the full record. When there is some freedom in setting education subsidies [i.e., any schedule such that $0 \leq S_s(i) \leq 1$] we are necessarily weakly better off because not intervening is feasible as well.

By (2.5) [which determines education choices], $T_s(i) = S_s(i) = 0, \forall i$ and the assumptions on the skill technology function, private agents choose $\{s_1, s_2\}^* > 0$. The first period government still optimizes over A , which in this case is determined by the optimality condition $[u'(c_1^y) + u'(c_2^y)]/2 = \beta(1+r)\tilde{\lambda}$.

To compare social welfare as of date zero, fix the $\{s_1, s_2, A\}^*$ allocation and suppose agents choose otherwise without government intervention. Given the infinite productivity of the first units of schooling and the shape of $\psi(\cdot)$, $s_i = 0$ is always worse than an interior value of schooling - namely s_i^* - when agents may freely choose (n_i, c_i^o) , unless s_i^* exhausts available resources in the first period [bringing $u'(c_i^y)$ close to infinity]. The fact that the first period government also sets A^* such that $[u'(c_1^y) + u'(c_2^y)]/2 = \beta(1+r)\tilde{\lambda}$ necessarily eliminates such scenario. Furthermore, the implementation of a second best in the second period improves social welfare vis-a-vis a second period *laissez faire*. Therefore, $W^y(c_1^y, c_2^y, c_1^o, c_2^o, n_1, n_2)$ is necessarily higher in limited record than in full record when the government lacks commitment and there is no education policy.

Finally, whatever the preferred tax schedule of an uncommitted government having to comply with limited record, it is necessarily feasible for the committed government also facing the limited record keeping restriction. Both are non-linear functions $T(Y)$, and for the same tax schedule, the first period preferred plan for education would be necessarily the same. Therefore, second period feasibility would verify equally and the incentive compatibility constraint (2.9) is exactly the same [see also section 5.3]. This means that the committed government achieves at least the same level of welfare as the uncommitted. It will generally do better because it optimizes also over incentives to invest in education in the first period when setting $T(Y)$. ■

Proof of Proposition 4:

The Lagrangian associated with programme (2.13) is

$$\begin{aligned} \mathcal{L} = & u(c_1^y) + u(c_2^y) + \beta V^{o,sep}(s_1, s_2, A) - \lambda_y^{sep}(c_1^y + c_2^y + s_1 + s_2 + A - k_0) \\ & + \mu_y^{sep} \left[u(c_1^y) - \beta v \left(\frac{Y_1}{\psi(\theta_1, s_1)} \right) - u(c_2^y) + \beta v \left(\frac{Y_2}{\psi(\theta_1, s_2)} \right) \right]. \end{aligned}$$

The first order necessary conditions include:

$$\begin{aligned} (c_1^y, c_2^y) : & u'(c_1^y)(1+\mu) = u'(c_2^y)(1-\mu) = \lambda_y^{sep} \\ (s_1) : & \beta \left[V_{s_1}^{o,sep} + \mu_y^{sep} v' \left(\frac{Y_1}{\psi(\theta_1, s_1)} \right) \frac{Y_1}{\psi(\theta_1, s_1)} \frac{\psi_s(\theta_1, s_1)}{\psi(\theta_1, s_1)} \right] = \lambda_y^{sep} \\ (s_2) : & \beta \left[V_{s_2}^{o,sep} - \mu_y^{sep} v' \left(\frac{Y_2}{\psi(\theta_1, s_2)} \right) \frac{1}{\psi(\theta_1, s_2)} \frac{\psi_s(\theta_1, s_2)}{\psi(\theta_1, s_2)} \right] = \lambda_y^{sep} \\ (A) : & \beta V_A^{o,sep} = \lambda_y^{sep} \\ & \mu_y^{sep}, \lambda_y^{sep} > 0 \end{aligned}$$

Applying the envelope theorem in (2.11) we obtain $V_{s_i}^{o,sep} = v' \left(\frac{Y_i}{\psi(\theta_i, s_i)} \right) \frac{Y_i}{\psi(\theta_i, s_i)} \frac{\psi_s(\theta_i, s_i)}{\psi(\theta_i, s_i)}, \forall i$ and $V_A^{o,sep} = u'(c^o)(1+r)$. Substituting and rearranging terms, and having into account that $\mu_y^{sep} < 1$ is necessary to meet condition (2.12), we obtain

$$\begin{aligned} u'(c_1^y) &< \beta u'(c^o)(1+r) < u'(c_2^y) \\ \frac{\beta v' \left(\frac{Y_1}{\psi(\theta_1, s_1)} \right) \frac{Y_1}{\psi(\theta_1, s_1)} \frac{\psi_s(\theta_1, s_1)}{\psi(\theta_1, s_1)}}{u'(c_1^y)} &= 1 \\ \frac{\beta v' \left(\frac{Y_2}{\psi(\theta_2, s_2)} \right) \frac{Y_2}{\psi(\theta_2, s_2)} \frac{\psi_s(\theta_2, s_2)}{\psi(\theta_2, s_2)}}{u'(c_2^y)} &= \frac{1 - \mu_y^{sep}}{1 - \mu_y^{sep} \frac{v' \left(\frac{Y_2}{\psi(\theta_1, s_2)} \right) \frac{1}{\psi(\theta_1, s_2)} \frac{\psi_s(\theta_1, s_2)}{\psi(\theta_1, s_2)}}}{\frac{v' \left(\frac{Y_2}{\psi(\theta_2, s_2)} \right) \frac{1}{\psi(\theta_2, s_2)} \frac{\psi_s(\theta_2, s_2)}{\psi(\theta_2, s_2)}}}. \end{aligned}$$

$U(c_1^y, c_1^o, n_1) > U(c_2^y, c_2^o, n_2)$ holds when (2.12) is met with equality. ■

Proof of Proposition 5: $V^{o,lr}(s_1, s_2, A)$ is the indirect utility derived from a programme equivalent to (2.23). With $\psi(\theta_1, s_1) > \psi(\theta_2, s_2)$, the first order conditions are also the same - (2.24) -, implying: $c_1^o > c_2^y$, $-MRS_{Y_1, c_1^o} = 1$ and $-MRS_{Y_2, c_2^o} < 1$. The expression for $-MRS_{Y_2, c_2^o}$ may be contrasted to (2.21). There we find the last term in the denominator concerning the impact of Y_2 on first period incentive constraints, which does not show up in this government's problem.

Using the envelope theorem we find the expressions for $V_{s_1}^{o,lr}$ and $V_{s_2}^{o,lr}$:

$$\begin{aligned} V_{s_1}^{o,lr} &= v' \left(\frac{Y_1}{\psi(\theta_1, s_1)} \right) \frac{\psi_s(\theta_1, s_1)}{\psi(\theta_1, s_1)} \frac{Y_1}{\psi(\theta_1, s_1)} \left\{ 1 + \mu_o^{lr} \left[1 - \frac{v' \left(\frac{Y_2}{\psi(\theta_1, s_1)} \right) \frac{Y_2}{\psi(\theta_1, s_1)}}{v' \left(\frac{Y_1}{\psi(\theta_1, s_1)} \right) \frac{Y_1}{\psi(\theta_1, s_1)}} \right] \right\} \\ V_{s_2}^{o,lr} &= v' \left(\frac{Y_2}{\psi(\theta_2, s_2)} \right) \frac{\psi_s(\theta_2, s_2)}{\psi(\theta_2, s_2)} \frac{Y_2}{\psi(\theta_2, s_2)}, \end{aligned}$$

where μ_o^{lr} is the multiplier to the second period incentive compatibility constraint. The first order conditions for the social optimization problem (2.16) with respect to consumption and schooling, after substituting for optimality in A , have exactly the same form as equations (2.18)-(2.20) [substituting \bar{x} by μ_o^{lr} and $\bar{\mu}$ by the multiplier of the incentive compatibility constraint in (2.16), μ_y^{lr}]. Finally $U(c_1^y, c_1^o, n_1) > U(c_2^y, c_2^o, n_2)$ always holds when both the first and second period incentive constraints hold with equality. ■

2.8 References

- Acemoglu, D., Golosov, M., Tsyvinski, A., (2008). Dynamic Mirrlees Taxation and Political Economy. *NBER working paper 12224*.
- Albanesi, S., Sleet, C., (2006). Dynamic Optimal Taxation with Private Information. *Review of Economic Studies* (73).
- Bester, H., Strausz, R., (2001). Contracting with Imperfect Commitment and the Revelation Principle: The Single Agent Case. *Econometrica* 69.
- Bester, H., Strausz, R., (2003). Contracting with Imperfect Commitment and Noisy Communication. *working paper, Free University Berlin*.
- Berliant, M., Ledyard, J. (2005). Optimal Dynamic Nonlinear Taxes with no Commitment. *mimeo*.
- Bisin, A., Rampini, A. A., (2006). Markets as Beneficial Constraints on the Government. *Journal of Public Economics*, vol. 90, issue 4-5.
- Boadway, R., Marceau, N., Marchand, M., (1996). Investment in Education and the Time Inconsistency of Redistributive Tax Policy. *Economica* Vol.63.
- Bovenberg, L., Jacobs, B., (2005). Redistribution and Education Subsidies are Siamese Twins. *Journal of Public Economics* (89).
- Bovenberg, L., Jacobs, B., (2008). Optimal Taxation of Human Capital and the Earnings Function. *CESifo working paper no.2162*.
- Bohacek, R., Kapicka, M., (2008). Optimal Human Capital Policies. *Journal of Monetary Economics*, vol. 55 (1).
- Brett, C., Weymark, J. A., (2008). Optimal Nonlinear Taxation of Income and Savings without Commitment. *Vanderbilt University working paper 08-W05*.
- Brett, C., Weymark, J. A., (2008a). The Impact of Changing Skill Levels on Optimal Nonlinear Income Taxes. *Journal of Public Economics*, 92.
- Carneiro, P., Heckman, J., (2003). Human Capital Policy. IZA Discussion Paper No. 821.
- Cohen, D., Michel, P., (1988). How Should Control Theory be Used to Calculate a Time-Consistent Government Policy? *Review of Economic Studies*, vol. 55, n.2.
- Dillén, M., Lundholm, M., (1996). Dynamic Income Taxation, Redistribution, and the Ratchet Effect. *Journal of Public Economics* 59.
- Fahri, E., Werning, I., (2008). The Political Economy of Nonlinear Capital Taxation. *working paper*.

- Fischer, S., (1980). Dynamic Inconsistency, Cooperation and the Benevolent Disassembling Government. *Journal of Economic Dynamics and Control*, 2.
- Freixas, X., Guesnerie, R., Tirole, J., (1985). Planning Under Incomplete Information and the Ratchet Effect. *Review of Economic Studies* 52.
- Gaube, T., (2007). Optimal Taxation of Each Year's Income. *Journal of Public Economic Theory*, 9 (1).
- Golosov, M., Tsyvinski, A., Werning, I. (2007). New Dynamic Public Finance: A user's Guide. In: *Acemoglu, D., Rogoff, K., Woodford, M. (Eds.), NBER Macroeconomics Annual 2006. Vol. 21. MIT Press.*
- Golosov, M., Kocherlakota, N., Tsyvinski, A., (2007). Optimal Indirect and Capital Taxation. *Review of Economic Studies*, 70.
- Konrad, K. (2001). Privacy, Time-Consistent Optimal Labor Income Taxation and Education Policy. *Journal of Public Economics* 79(3).
- Kapicka, M., (2006). Optimal Income Taxation with Human Capital Accumulation and Limited Record Keeping. *Review of Economic Dynamics* (9).
- Kocherlakota, N., (2005). Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation. *Econometrica* vol. 73 (5).
- Krause, A., (2008). Optimal Nonlinear Taxation with Learning-by-Doing. *mimeo, University of York*.
- Maldonado, D., (2007). The Design of Optimal Education Policies when Individuals Differ in Inherited Wealth and Ability. *Documento de Trabajo No. 21, Universidad del Rosario*.
- Mirrlees, J., (1971). An Exploration in the Theory of Optimal Income Taxation. *Review of Economic Studies* (38).
- Ortigueira, S., (2006). Markov-perfect optimal Taxation. *Review of Economic Dynamics*, vol. 9.
- Pereira, J., (2008). Commitment and Redistribution with Endogenous Skills. *mimeo, European University Institute*.
- Reis, C. (2005). Education and Optimal Taxation. *mimeo, MIT*.
- Roberts, K., (1984). The Theoretical Limits to Redistribution. *Review of Economic Studies* 51.
- Simula, L. (2007), Optimality Conditions and Comparative Static Properties of Nonlinear Income Taxes Revisited, *Paris School of Economics*, WP2007-15.

Stiglitz, J. (1982). Self-Selection and Pareto Efficient Taxation. *Journal of Public Economics*, vol 17(2).

Werning, I., (2007). Optimal Fiscal Policy With Redistribution. *Quarterly Journal of Economics* (August 2007).

CHAPTER 3

MARKOV-PERFECT OPTIMAL FISCAL POLICY: THE CASE OF UNBALANCED BUDGETS

3.1 Introduction

¹In most developed economies, income tax rates and government debt levels are positive and sizable. In the US, effective income tax rates have been in the order of 20%, and outstanding public debt represents about 60% of GDP. The main question we pose in this paper is: can these numbers be accounted for as the outcome of a government's welfare maximization program in a neoclassical economy? To address this question we adopt the standard framework in the literature of optimal fiscal policy, and drop the assumption of government's full commitment to future policies. Instead, we assume that the government has no access to commitment devices nor to reputation mechanisms, and, therefore, we restrict our attention to Markovian optimal policies. Our answer to the above question is in the affirmative, provided this is the policy expected by households and all successive governments.

The observation of positive income taxes and, especially, of positive levels of public debt has been at odds with most neoclassical theories of optimal fiscal policy. Indeed, the now classical result by Chamley (1986) and Judd (1985) establishes that a committed government will not use distortionary taxation in the long run. The optimal policy set by such a government involves high taxation in the short run in order to build up enough assets to finance future government expenditure, so that distortionary taxation can be disposed of in the long run. In economies with government's full commitment, this result has been proved to be robust to a number of non-trivial departures from the standard framework.

In this paper, we study a neoclassical economy populated by infinitely-lived consumers, competitive firms operating a constant-returns-to-scale production technology, and a benevolent government. The government makes sequential decisions on the provision of a valued public good, on income taxation and the issue of public debt. We characterize and compute Markov-perfect optimal fiscal policy in this economy with two payoff-relevant state variables: physical capital and public debt. Other than imposing differentiable strategies, we do not

¹This chapter is co-authored with Salvador Ortigueira.

restrict further the definition of Markov perfection. Hence, we look at all Markov-perfect equilibria of the infinite-horizon economy, including those which are the limit of equilibria of the finite-horizon economy and those which only emerge with infinite horizons.

The main contribution of our analysis can be summarized as follows. In the class of economies outlined above, optimal fiscal policy in a Markov-perfect equilibrium is not uniquely determined. We find two stable, steady-state equilibria: in one of them income taxes and debt are positive, implying positive distortions to both the consumption/savings and the private/public consumption margins. In the second equilibrium taxes are zero and debt is negative, implying no distortions to these two margins. Moreover, in a calibrated version of the model that matches key US observations, we show that a 20% income tax rate and a debt-GDP ratio of 60% emerge as the long-run optimal fiscal policy in the equilibrium with positive distortions. We prove that convergence to either of the two long-run equilibria is not pinned down by initial conditions but by expectations on government policy. That is, Markov-perfect optimal fiscal policy is driven by expectations.

In the economy of our model, the multiplicity of expectation-driven Markov-perfect equilibria does not arise if the government is required to balance its budget on a period-by-period basis, in which case income taxation becomes the only source of government revenue. It is only when governments are allowed to run unbalanced budgets and, therefore, to spread the burden of financing the provision of the public good that expectations play a role in the determination of optimal fiscal policy. Thus, expectations that all future governments will dispose of distortionary taxation if given enough assets to finance the provision of the public good, will render such a policy optimal. On the other hand, expectations that all governments will issue debt in order to pass on part of the burden of financing current expenditure to the next government will lead to an optimal policy with income taxation, issues of debt and, consequently, positive wedges. As we show below, the existence of this latter equilibrium hinges on the assumption that the economy runs for an infinite number of periods, and therefore there is no last government unable to pass on the burden.² A feature common to the two equilibria is that governments use public debt to reduce long-run tax distortions, as compared to the economy without debt.

In economies without capital, the existence of two steady-state Markov-perfect equilibria has been recently shown by a number of authors. In these economies, however, equilibrium dynamics are drastically different from what we find in our economy with capital and debt.

²The expectational multiplicity of the equilibrium in our economy is thus of a different nature to that found by Calvo (1988) in a two-period economy with public debt and costly debt repudiation. This author shows the existence of two expectation-driven equilibria: a “good” Pareto-efficient equilibrium in which there is no debt repudiation, and a “bad” Pareto-inefficient equilibrium where debt is partially repudiated.

Martin (2006) and Díaz-Giménez, Giovannetti, Marimon and Teles (2006), study optimal monetary policy in economies with debt and find two steady-state debt levels, only one of which is stable. These authors show that the two steady states are generated by the same policy function and, therefore, the Markov-perfect equilibrium is unique. Krusell, Martin and Ríos-Rull (2006) study optimal debt policy in a model with exogenous government expenditure, labor taxation and no capital. In their economy, the interior steady-state equilibrium with positive distortions is unstable. The authors show that the equilibrium contains a large, countable set of long-run debt levels. Initial conditions pin down the element in this set to which the economy converges in a maximum of two periods.

Our paper builds on a large body of literature dealing with optimal fiscal policy in environments with no commitment. Thus, Markov-perfect optimal taxation in economies without public debt has been first studied by Klein and Ríos-Rull (2003), Klein, Krusell and Ríos-Rull (2006) and then by Ortigueira (2006). Indeed, our paper is an extension of the framework presented in Klein, Krusell and Ríos-Rull (2006) to include public debt.

Our paper is also related to the work of Song, Storesletten and Zilibotti (2007) who study optimal fiscal policy in economies where subsequent generations of agents (young and old) vote on policy. These authors focus on the Markov-perfect political equilibrium in an economy without physical capital and find that the long-run level of debt depends crucially on the distortions brought about by labor taxation. When these distortions are large enough debt converges to an interior value, otherwise debt accumulation depletes the economy. Debortoli and Nunes (2007) assume political disagreement (i.e. policymakers have different preferences on the type of public good that should be provided) as in Alesina and Tabellini (1990) and Persson and Svensson (1989) and study the evolution of public debt under different degrees of commitment. Abstracting from physical capital, the authors show that it is political disagreement what explains positive values of long-run debt.

A different body of literature has developed after the paper by Lucas and Stokey (1983), who study the role of public debt as a substitute for commitment in Ramsey economies without capital. They show that the Ramsey policy is consistent if governments can commit to inherited debt contracts. Specifically, they show that future governments will comply with the fiscal plans chosen today if the current government delegates rich enough state-contingent multiple-period debt contracts. Later, Persson, Persson and Svensson (1988) extend this line of research to monetary policy inconsistency. Finally, Aiyagari, Marcet, Sargent and Sepúlveda (2002) modify the Stokey and Lucas (1993) model by dropping the complete markets assumption, which introduces a history dependence on the debt path, as opposed to a contingency to future states. They show that when there are no exogenous bounds on debt the

Ramsey planner in their economy lets public debt converge to a negative level.

The chapter is organized as follows. Section 3.2 presents the model, characterizes Markov-perfect equilibria and shows the existence of the two steady states. In Section 3.3 we parameterize and calibrate our model economy and compute Markov-perfect equilibrium. We also compare Markovian policies with those arising in the efficient and the Ramsey equilibrium. Section 3.4 presents a description of our numerical algorithm. Section 3.5 concludes and Section 3.6 contains the Appendix.

3.2 The Model

Our framework is the standard, non-stochastic neoclassical model of capital accumulation, extended to include a benevolent government that provides a valued public good. In order to finance the provision of such public good the government can levy a tax on household's income and issue public debt. Thus, fiscal policy in each period consists of the amount of the public good provided, G_t , the tax rate on income, τ_t , and the issue of public debt, B_{t+1} , which matures in period $t + 1$.

We begin by describing the problem solved by each agent in this economy. We then characterize the fiscal policy set by the benevolent government lacking the ability to commit to future policies. In order to help compare our results with the case of full commitment, we also present a brief review of fiscal policy in the Ramsey equilibrium.

3.2.1 Households

There is a continuum of homogeneous households with measure one. Each household supplies one unit of labor and chooses consumption and savings in order to maximize lifetime utility, subject to a budget constraint and initial endowments of physical capital and public debt,

$$\max_{\{c_t, k_{t+1}, b_{t+1}\}} \sum_{t=0}^{\infty} \beta^t U(c_t, G_t), \quad (3.1)$$

s.t.

$$c_t + k_{t+1} + b_{t+1} = k_t + b_t + (1 - \tau_t) [w_t + (r_t - \delta) k_t + q_t b_t] \quad \forall t \quad (3.2)$$

$k_0 > 0$ and b_0 given,

where small letters are used to denote individual variables and capital letters to denote economy-wide values. Function $U(\cdot)$ in equation (3.1) is the instantaneous utility function, which depends on the consumption of a private good, c_t , and the consumption of a public

good, G_t . $U(\cdot)$ is assumed to be continuously differentiable, increasing and concave; and $0 < \beta < 1$ is the discount factor. Labor is supplied inelastically at a real wage rate w_t . Household's asset holdings are made up of physical capital, k_t , which is rented to firms at the rate r_t , and government's bonds, b_t , which bear an interest denoted by q_t . Physical capital depreciates at a rate denoted by $0 < \delta < 1$. Household's total income, net of capital depreciation, is taxed at the rate τ_t . If the government is a net lender to the private sector—i.e., the household borrows from the government, $b_t < 0$ —taxable income is net of interest payments.

3.2.2 Firms

Firms are competitive and produce an aggregate good with a neoclassical production technology. Total production is given by,

$$Y_t = F(K_t, L_t) = F(K_t, 1) = f(K_t) \quad \forall t, \quad (3.3)$$

where K_t denotes the aggregate or economy-wide stock of capital and we have $f'(K_t) > 0$, $f''(K_t) < 0$. First-order conditions to profits maximization imply the typical demand and zero-profits equations,

$$r_t = f_K(K_t) \quad (3.4)$$

$$w_t = f(K_t) - r_t K_t. \quad (3.5)$$

3.2.3 Government

Government's fiscal policy involves the setting of both the provision of the public good and its financing through taxes and debt. The government is benevolent in the sense that it seeks to maximize households' lifetime utility, (3.1), subject to its budget constraint, to a feasibility restriction, and to private sector's first-order conditions. In addition, government's policies may be conditioned by its lack of commitment. The budget constraint of the government is,

$$G_t + (1 + q_t)B_t = B_{t+1} + \tau_t [w_t + (r_t - \delta) K_t + q_t B_t], \quad (3.6)$$

where the right-hand side of equation (3.6) represents government's revenues, which are made up of the issue of debt, B_{t+1} , plus revenues from income taxation. The left-hand side is government's total expenditure, including the provision of the public good, the repaying of outstanding public debt and financial expenses.

3.2.4 Ramsey Optimal Fiscal Policy

This Section presents a brief review of the Ramsey fiscal policy in our model economy. In a Ramsey equilibrium, the benevolent government is assumed to have full commitment to future policies, and, thus, it can credibly announce the whole sequence of public expenditure, income taxes and issues of debt from the first period onwards. This allows the government to anticipate the response of the private sector to its fiscal policy. Hence, the problem of the government in the Ramsey equilibrium is to choose sequences for taxes and public debt so that the competitive equilibrium maximizes social welfare [equation (1.1)].

Proposition 1 below presents the optimal fiscal policy in the steady-state Ramsey equilibrium for our economy. Since the result in Proposition 1 is well known in the literature of optimal fiscal policy we only provide a sketch of the proof (see the Appendix).

Proposition 1: *In the steady-state Ramsey equilibrium the tax rate on income is zero and the government holds positive assets, i.e. $B < 0$.*

3.2.5 Markov-Perfect Optimal Fiscal Policy

In this Section we drop the assumption of government's full commitment to future policies and study time-consistent optimal policies. More specifically, we will focus on differentiable Markov-perfect equilibria of this economy populated by a continuum of households and a government that acts sequentially, foreseeing its future behavior when choosing current levels of the public good, income taxes and the issue of debt. The restriction to differentiable Markov-perfect policies is justified by our use of calculus in the characterization of the equilibria. A further remark on the assumption of differentiability will follow below.

Following the literature on Markovian policies, we assume that the government — although unable to commit to future policies —, does commit to honoring the tax rate it announces for the current period, and to repaying outstanding debt obligations. The commitment to current taxes implies an intra-period timing of actions that grants the government leadership in the setting of the tax rate. That is, at the beginning of period t , the time- t government sets the tax rate for the period; next, once that choice is publicly known, consumers choose consumption/savings and the composition of their portfolios, and the government chooses the provision of the public good (or equivalently, the issue of debt). Governments are thus (intra-period) *Stackelberg players* and can therefore anticipate the effects of current taxation on household's decisions.

In sum, we assume that the time- t government has intra-period commitment to time- t taxes but not to debt issues. In our opinion, this fits well the timing of actions in real economies, where, typically, governments make decisions on taxes at discrete times but issue

debt continuously. [For a discussion on the effects of the timing of actions on Markovian policies see Ortigueira (2006).]

The optimization problem of a typical household

The household chooses (i) how much to consume and save; and (ii) how to allocate savings between physical capital and public debt. At the time the household makes these decisions the tax rate for the period is known, but it must foresee both the current government's debt policy and future governments' fiscal policy.

Hence, the problem of a household that holds k and b of the physical and government assets, respectively, that has to pay taxes on current income at rate τ , that expects the current and future governments to issue new debt according to the policy $\psi_B : (K \times B \times \tau) \rightarrow B'$, and expects future governments to set taxes according to the policy $\psi_\tau : (K \times B) \rightarrow \tau$, can be written as,

$$\begin{aligned} v(k, b, K, B; \tau) &= \max_{c, k', b'} \{U(c, G) + \beta \tilde{v}(k', b', K', B')\} \\ &\text{s.t.} \\ c + k' + b' &= k + b + (1 - \tau) [w(K) + [r(K) - \delta] k + q(K)b], \end{aligned} \quad (3.7)$$

where $\tilde{v}(k', b', K', B')$ is the continuation value as foreseen by the household. $\omega(K)$, $r(K)$ and $q(K)$ are pricing functions. The economy-wide stock of physical capital is expected to evolve according to the law $K' = \mathcal{H}(K, B, \tau)$, say. By using the assumption of a representative agent, i.e., $k = K$ and $b = B$, and the government's budget constraint, it follows from the above maximization problem that the consumption function in a competitive equilibrium—where today's tax rate is τ , future taxes are set according to policy ψ_τ and current and futures issues of debt are set according to policy ψ_B —can be expressed in terms of K, B and τ , say $C(K, B, \tau)$, and must satisfy the following Euler equation,

$$U_c(C(K, B, \tau), G) = \beta U_c(C(K', B', \tau'), G') [1 + (1 - \tau') (f_K(K') - \delta)], \quad (3.8)$$

where $B' = \psi_B(K, B, \tau)$ and $\tau' = \psi_\tau(K', B')$. In equilibrium K' is given by,

$$K' = K + B + (1 - \tau) [f(K) - \delta K + q(K)B] - C(K, B, \tau) - B', \quad (3.9)$$

where G and G' are given by the time- t and time- $(t + 1)$ governments' budget constraints, respectively. Finally, pricing functions $\omega(K)$ and $r(K)$ are given by (3.4) and (3.5), and $q(K)$ must satisfy the non-arbitrage condition between the two assets,

$$q(K) = f_K(K) - \delta. \quad (3.10)$$

In equilibrium, capital and debt yield the same return, meaning that q is independent of B . The fact that the interest rate on public debt is independent of B implies an important departure from economies without physical capital. We will comment further on this issue below.

Equation (3.8) has the usual interpretation: the marginal utility of consumption equals the present value of the last unit of income devoted to savings. Since physical capital and debt yield the same return in equilibrium, the supply of public debt determines the composition of the household's portfolio. This implies a one-to-one crowding out of investment in capital by public debt. Taxation, on the other hand, affects disposable income and the level of consumption, and thus translates into a non-one-to-one crowding out of capital investment. The problem of the government is shown next.

The problem of the government

As explained above, the government's lack of commitment to future policies and our focus on Markov-perfect equilibria allows us to think of the government as a sequence of governments, one for each time period. Thus the time- t government sets the tax rate for the period and issues new debt foreseeing the fiscal policy to be set by successive governments. Following the timing of actions established above, the time- t government is an intra-period *Stackelberg player* in our economy: At the beginning of the period, it chooses the income tax rate for that period taking into account the effect of τ on the level of consumption, as given by the consumption function, $C(K, B, \tau)$, that solves (3.8) and (3.9). In a second stage, the government sets the issue of debt. The problem of the government is thus solved backwards. Given the initial choice for taxes, the issue of debt is the solution to,

$$\begin{aligned} V(K, B, \tau) &= \max_{B'} \left\{ U(C(K, B, \tau), G) + \beta \tilde{V}(K', B') \right\} \\ \text{s.t.} \quad & K' = (1 - \delta)K + f(K) - C(K, B, \tau) - G \\ & G = \tau [f(K) - \delta K + q(K)B] + B' - [1 + q(K)] B, \\ & \text{and equation (3.10),} \end{aligned} \tag{3.11}$$

where $V(K, B, \tau)$ is the value to the time- t government that has set the tax rate at τ and foresees the fiscal policy to be set by future governments. $\tilde{V}(K', B')$ is next-period value as foreseen by the time- t government. The issue of debt that solves this problem can thus be written as $B'(K, B, \tau)$. Therefore, the tax rate set by the time- t government is the solution

to,

$$\begin{aligned}
 W(K, B) &= \max_{\tau} \left\{ U(C(K, B, \tau), G) + \beta \tilde{V}(K', B'(K, B, \tau)) \right\} \\
 \text{s.t.} \quad & \\
 K' &= (1 - \delta)K + f(K) - C(K, B, \tau) - G \\
 G &= \tau [f(K) - \delta K + q(K)B] + B'(K, B, \tau) - [1 + q(K)]B, \\
 &\text{and equation (3.10).}
 \end{aligned} \tag{3.12}$$

The following proposition characterizes the fiscal policy set by the time- t government.

Proposition 2: *The tax and debt policy that solves the government's problem is the solution to the following Generalized Euler Equations:*

$$\frac{U_c C_{\tau} + U_G G_{\tau}}{G_{\tau} + C_{\tau}} = \beta \left[U'_c C'_{K'} + U'_{G'} G'_{K'} + \frac{U'_c C'_{\tau'} + U'_{G'} G'_{\tau'}}{G'_{\tau'} + C'_{\tau'}} (f'_{K'} + 1 - \delta - C'_{K'} - G'_{K'}) \right] \tag{3.13}$$

and

$$\frac{U_c C_{\tau} + U_G G_{\tau}}{G_{\tau} + C_{\tau}} = U_G G_{B'} + \beta \left[U'_c C'_{B'} + U'_{G'} G'_{B'} - \frac{U'_c C'_{\tau'} + U'_{G'} G'_{\tau'}}{G'_{\tau'} + C'_{\tau'}} (C'_{B'} + G'_{B'}) \right]. \tag{3.14}$$

Proof: See the Appendix.

Some comments on notation are in order. Function arguments in equations (3.13) and (3.14) have been omitted for expositional clarity. Subscripts denote the variable with respect to which the derivative is taken. A prime in a variable indicates next-period values, and a prime in a function indicates it is evaluated at next-period variables. Finally, G_{τ} and G_B denote the derivatives of G with respect to τ and B , respectively, holding B' constant.

Before providing an interpretation of the two Generalized Euler Equations presented in Proposition 2, we offer the following definition. A Markov-perfect equilibrium in our economy can be loosely defined as:

Definition: *A Markov-perfect equilibrium is a quadruplet of functions $C(K, B, \tau)$, $\psi_B(K, B, \tau)$, $\psi_{\tau}(K, B)$ and $W(K, B)$, such that:*

- (i) *Given ψ_B and ψ_{τ} , $C(K, B, \tau)$ solves the household's maximization problem.*
- (ii) *Given $C(K, B, \tau)$, ψ_B and ψ_{τ} solve the government's maximization problem. That is, $B' = \psi_B(K, B, \tau)$ and $\tau = \psi_{\tau}(K, B)$.*

(iii) $W(K, B)$ is the value function of the government.

The two Generalized Euler Equations, (3.13) and (3.14), which characterize Markov-perfect taxation and debt policies, respectively, have the following interpretation. Equation (3.13) establishes that the tax rate has to equate the marginal value of taxation to the marginal value of investing in physical capital. Equation (3.14) establishes that the issue of debt has to equate the marginal value of issuing debt to the marginal value of investing in physical capital (and consequently to the marginal value of taxation). In a Markov-perfect equilibrium, the government is indifferent between using taxes or debt to finance the provision of the public good. Both equations involve only wedges between today and tomorrow, as posterior wedges are implicitly handled optimally by an envelope argument. Consecutive governments, however, disagree on how much to tax tomorrow [the time- $(t+1)$ government does not internalize the distortionary effects of its policy on time- t investment]. The current government thus takes into account the effect of its policy on tomorrow's initial conditions, K' and B' , in order to help compensate for that disagreement. Following this reasoning, one may interpret the different terms in (3.13) and (3.14) as follows.

The left-hand side of equation (3.13) is today's marginal utility of taxation *per* unit of savings crowded out. The numerator of this expression is the change in utility from a marginal increase in the tax rate, which is made up of the change in utility from the private good, $U_c C_\tau$, plus the change in utility from the public good, $U_G G_\tau$. The denominator is the amount of savings crowded out, or, equivalently, the change in consumption of the public and private good brought about by the increase in the tax rate.

The right-hand side of equation (3.13) is the marginal utility of investing in physical capital. An extra unit of investment today yields an increase in resources tomorrow by $f'_{K'} + 1 - \delta$. The breakdown of the value of these resources is: (i) $C'_{K'}$ of them are consumed as private good, yielding a value of $U'_c C'_{K'}$; (ii) $G'_{K'}$ corresponds to the increase in the provision of the public good obtained from the increase in the tax base, which yields a value of $U'_G G'_{K'}$; (iii) the remaining $f'_{K'} + 1 - \delta - C'_{K'} - G'_{K'}$ are taxed away, and the marginal value is the left-hand side of equation (3.13), updated one period ahead. Hence, the right-hand side of (3.13) results from adding up all these values and discounting.

Equation (3.14) is a non-arbitrage condition between taxation and public debt, and its interpretation is equally straightforward. The right-hand side is the value of issuing an extra unit of government debt today. The first term on the right-hand side is the value of today's extra public good financed with the increase in government debt. The second term is the present value of the implied changes in tomorrow's consumption of the private and public good, $C'_{B'}$ and $G'_{B'}$, respectively. Besides the direct effects on tomorrow's utility, these

changes have an effect on tomorrow's taxation, which must be valued using the marginal utility of taxation. Equation (3.14) establishes that the value of issuing debt must equal the value of taxation (the left-hand side of the equation).

A re-arrangement of equation (3.14) offers an alternative interpretation of the non-arbitrage condition between taxes and bonds in terms of two wedges, $U_c - U_G$ and $U'_{c'} - U'_{G'}$. Such a re-arrangement yields,

$$(U_c - U_G) \frac{C_\tau}{G_\tau + C_\tau} + \beta \left\{ (U'_{c'} - U'_{G'}) \left(G'_{B'} + \frac{G'_{\tau'}}{G'_{\tau'} + C'_{\tau'}} K''_{B'} \right) \right\} = 0. \quad (3.15)$$

Equation (2.15) says that the value of using debt instead of taxes to finance the last unit of public expenditure equals zero in a Markov-perfect equilibrium. The first term is the net change in utility today of using debt instead of taxes per unit of forgone savings. The second term captures the change in future distortions induced by the extra unit of public debt. The way the current government trades off these two wedges when choosing B' depends on expectations on future government policy. As will become clearer below, there is an equilibrium policy which renders a non-zero wedge $U_c - U_G$ in the long run.

3.2.5.1 Steady-State Markov-Perfect Equilibrium

A *steady-state Markov-perfect equilibrium* is defined as a list of infinite sequences for quantities $\{C_t\}$, $\{K_t\}$, fiscal variables, $\{G_t\}$, $\{\tau_t\}$, $\{B_t\}$ and prices $\{\omega_t\}$, $\{r_t\}$, and $\{q_t\}$ such that they are generated by a Markov-perfect equilibrium, and its values do not change over time, i.e. $K_{t+1} = K_t$, $B_{t+1} = B_t$, $\tau_{t+1} = \tau_t$ for all t , and the same is true for consumption and prices.

In this subsection we offer some insights on the steady-state Markov-perfect equilibrium of our model economy, and prove three propositions. A first insight is related to the existence of two different steady-state Markov-perfect equilibria. Evaluating equation (2.15) at a steady-state Markov-perfect equilibrium yields,

$$(U_c - U_G) \left\{ \frac{C_\tau}{G_\tau + C_\tau} + \beta \left(G_B + \frac{G_\tau}{G_\tau + C_\tau} K'_B \right) \right\} = 0. \quad (3.16)$$

This equation suggests that there may be two different taxation and debt policies consistent with the existence of a steady-state Markov-perfect equilibrium. The first one corresponds to the policy prescribed by the long-run Ramsey equilibrium. As shown in Proposition 1, the Ramsey equilibrium prescribes zero income taxes and positive government asset holdings in the steady state. The provision of the public good is financed entirely from the

returns on government's assets, and therefore, $U_c = U_G$. The next proposition proves that this policy is a Markov-perfect equilibrium.

Proposition 3: *The steady-state Ramsey equilibrium is a Markov-perfect equilibrium.*

Proof: See Appendix.

In a related paper, Azzimonti-Renzo, Sarte and Soares (2006) study a model with differentiated taxes on capital and labor, and exogenous government expenditure. Within their framework, the authors find a Markov-perfect equilibrium which yields zero labor taxes from all initial conditions, K and B , and zero capital taxes from next-period onwards. As confirmed by our numerical computations, this result also holds in our model economy: when there are no exogenous bounds on income taxation, there exists a Markov-perfect equilibrium in which income taxes are zero after one period, and government assets converge to the long-run Ramsey value. Furthermore, for some initial conditions the initial income tax is negative, which amounts to a subsidy to households.

The second taxation and debt policy consistent with a steady-state Markov-perfect equilibrium involves positive income taxes and the issuing of government's bonds. Under this policy $U_c \neq U_G$, and the second term on the left-hand side of equation (3.16) is zero. The next proposition presents an important feature of the steady-state Markov-perfect equilibrium with positive taxation.

Proposition 4: *Along a steady-state Markov-perfect equilibrium with positive distortions, government bonds are not net wealth, i.e., $C_B(K^*, B^*, \tau^*) = 0$.*

Proof: See the Appendix.

Even though we do not have a formal proof establishing that the maximum number of stable, interior steady-state equilibria is two, our numerical computations lead us to be confident that this is the case. Our exploration of different subsets of the state space produced only a steady-state equilibrium with positive taxation and public debt.

The existence of two stable, steady-state equilibria raises a question concerning equilibrium dynamics from initial values K_0 and B_0 . Proposition 5 below proves that the government's policy rules generating the two steady states are different, which implies that steady-state multiplicity is expectational. Therefore, given initial conditions K_0 and B_0 , expectations on government policy determine equilibrium dynamics and convergence to one of the two long-run equilibria. The basic idea of the proof relies on the fact that the steady-state equilibrium with positive distortions is not the limit of the finite-horizon economy's Markov-perfect equilibrium as the time horizon goes to infinity. Actually, we show that the steady-state equilibrium with no distortions is the only limit of the finite-horizon equilibrium.

Proposition 5: *The two steady-state Markov-perfect equilibria are not associated with the*

same pair of decision rules ψ_τ and ψ_B . Hence, given K_0 and B_0 , the Markov-perfect equilibrium is (globally) indeterminate.

Proof: See the Appendix.

The next section presents a numerical analysis of the global dynamic properties of the steady-state Markov-perfect equilibrium with positive distortions.

3.3 The Markov-Perfect Equilibrium in a Calibrated Economy

In this section we parameterize our model economy, set values to its parameters and compute Markov-perfect equilibria. A special attention will be devoted to the presentation of the Markov-perfect equilibrium rendering distortary taxation and positive debt in the long run. We also compare Markov-perfect equilibria to the efficient solution (lump-sum taxation). Finally, we solve for the Markov-perfect equilibrium under balanced budgets and compare the results to the Ramsey and efficient solutions. A detailed explanation of our computational approach can be found in the next section.

The instantaneous utility function is assumed to be of the CES form in the composite good $c_t G_t^\theta$, that is,

$$U(c, G) = \frac{(c G^\theta)^{1-\sigma} - 1}{1-\sigma}, \quad (3.17)$$

where $0 < \theta < 1$, and $1/\sigma$ denotes the elasticity of intertemporal substitution of the composite good. The functional form for the production technology is the standard Cobb-Douglas function, with α denoting the capital's share of income, i.e.,

$$f(K) = AK^\alpha, \quad A > 0. \quad (3.18)$$

Parameter values are set as follows. The constant in the production function, A , and the inverse of the elasticity of intertemporal substitution, σ , are both set equal to one. The value of α is set at 0.36, which is the capital's share of income in the US economy; the depreciation rate of capital is set at 0.09, which is a standard value in macroeconomic models; β is set at 0.96, and θ is 0.2 so that the public-to-private consumption ratio falls within the range 15 – 30% for all equilibrium concepts mentioned above. These parameter values are in line with those in Klein, Ríos-Rull and Krusell (2006), Ortigueira (2006) and many others.

We start by presenting steady-state values for macroeconomic aggregates and fiscal policy under three equilibrium concepts —namely, the equilibrium with lump-sum taxes, the Ramsey equilibrium and the Markov-perfect equilibrium. Table 1 below presents these steady-state values.

Table 1
Steady-State Equilibria

	Efficient	Markov-perfect	
		No wedges (Ramsey)	Positive wedges
Y	1.7608	1.7608	1.6934
K	4.8144	4.8144	4.3201
C	1.1063	1.1063	1.1017
G	0.2213	0.2213	0.2032
G/C	0.2	0.2	0.1844
τ	indet.	0	0.1905
B/Y	indet.	-3.015	0.5639
W	-5.0157	-5.0157	-5.5525

Notes: Steady-state values and policy for the efficient, Ramsey and Markov-perfect equilibria.

The first column in Table 1 shows the efficient equilibrium when lump-sum taxation is available. The second and third columns show the two long-run Markov-perfect equilibria. The one in the second column is the equilibrium with zero wedges, which corresponds to the Ramsey equilibrium. In this equilibrium, the government does not distort long-run investment and sets income taxes equal to zero. Public expenditure is financed entirely from the income generated by the assets owned by the government. That is, negative public debt (positive asset holdings) is the only source of income for the government in this steady-state equilibrium. In the calibrated economy the value of the assets held by the government is larger than the assets held by the private sector, and more than three times the value of output.

The steady-state Markov-perfect equilibrium with positive wedges is shown in the third column of Table 1. In this equilibrium income is taxed at a rate of 19.05% and the debt-GDP ratio is 56.39%. These numbers fall well within the range of observed values in the U.S. and in most developed economies.

In our economy with physical capital accumulation and public debt, the configuration of Markovian equilibria differs drastically from that of the economy without capital, studied by Krusell, Martin and Ríos-Rull (2006). Contrary to their results, our steady-state equilibrium with positive distortions is stable. I.e. there is a Markov-perfect equilibrium whose time paths converge to this steady state, both for economies starting with debt levels below and above the steady-state value. An explanation as to why the presence of physical capital makes such an important impact in terms of equilibrium dynamics must be found in the

determination of the equilibrium interest rate on public debt. In our economy with capital, the interest rate is independent of the level of outstanding debt and is pinned down by the stock of capital. Thus, the current government can only affect next period's interest rate through the stock of capital. This is in contrast with the economy without capital where the government nails down next period's interest rate when setting today's debt issuance.

Figures 1 to 7 below display equilibrium dynamics converging to the Markov-perfect equilibrium with long-run distortions.³ (Details on our method to compute Markov-perfect equilibria can be seen below and in the next section.) Figures 1 to 3 show government's optimal fiscal policy along the Markov-perfect equilibrium converging to the steady state in the last column of Table 1. The optimal income tax, as a function of K and B , is shown in Figure 1. The tax rate increases both with capital and debt. Figure 2 shows government's debt policy. The issue of debt decreases sharply with capital, indicating that capital-rich economies rely relatively less on public debt to finance government. Figure 3 shows public expenditure as a function of K and B . The private-good consumption function is displayed in Figure 4.

The stability of the steady-state is shown in Figures 5 to 7. Net investment in physical capital, $K' - K$, is presented in Figure 5. In Figure 6 we plot the change in the level of outstanding debt, $B' - B$. Finally, Figure 7 presents the two loci, $K' = K$ and $B' = B$. The point in which these two loci intersect corresponds to the steady-state values for K and B . The arrows indicate the direction of the trajectories starting in the different regions of the state space.

It should be noted that the Markov-perfect equilibrium shown in Figures 1 to 7 has been computed using a global method. It becomes evident from a simple inspection of the Euler equation and the two Generalized Euler equations that the standard method of linearizing around steady-state values cannot be applied in our setting. Indeed, the equilibrium must be computed without prior knowledge of steady-state values. Thus, the subset of the state space must be changed in a trial-and-error process until it contains the steady-state equilibrium. Before moving on to the next section where we explain our computation strategy, we draw attention to Figures 8 to 11 below. Figures 8, 9 and 10 plot relative residuals in the Euler equation and the two Generalized Euler equations, respectively. Figure 11 shows relative residuals in the Bellman equation. It should be noted that the errors are very small, less than 0.001 of 1 per cent. In addition to this, the errors nearly satisfy the equioscillation property (the sign of the errors alternates between positive and negative), and show almost

³Equilibrium dynamics converging to the steady state without distortions are not presented here as they are well known from the literature on Ramsey optimal policy. (See the paragraph following Proposition 3.)

equal amplitude throughout the considered subset of the state space. All these properties of the errors indicate that our approximations are close to being optimal, in the sense that there are no better polynomials to approximate the unknown functions.

3.3.1 The Role of Debt in the Markov-Perfect Equilibrium

We now assess the role of public debt in economies without commitment. We present steady-state values for the three equilibrium definitions when governments are restricted to run balanced budgets. Table 2 below shows these values.

Table 2

Steady-State Equilibria with balanced budgets

	Efficient	Ramsey	Markov-perfect
Y	1.7608	1.7011	1.6710
K	4.8144	4.3742	4.1632
C	1.1063	1.0895	0.9911
G	0.2213	0.2179	0.3052
G/C	0.2	0.2	0.3079
τ	$T = G$	0.1666	0.2354
W	-5.0157	-5.4756	-6.1565

Notes: Steady-state values and policy for the efficient, Ramsey and Markov-perfect equilibria in the economy with balanced budgets.

As discussed above, the multiplicity of Markov-perfect equilibria does not hold in the economy with balanced budgets. The unique long-run Markov-perfect equilibrium in the economy without debt is shown in the third column of Table 2. This Markov-perfect equilibrium yields higher income taxes, underconsumption of the private good and overconsumption of the public good. Actually, the G/C ratio in the Markovian economy is 50% larger than in the efficient equilibrium, and the capital stock is 15% lower. The proneness of Markovian governments to overspend and overtax in the economy without debt is especially acute due to: (i) their lack of ability to internalize the distortionary effects of current taxation on past investment, and (ii) their leadership to set taxes before households choose consumption, which allows them to anticipate the response of current consumption to taxes, and then diminishing the perceived crowding out of physical investment.

Public debt plays a key role curbing the tendency of Markovian governments to overtax in the long run. In the steady-state Markov-perfect equilibrium with distortionary taxation the G/C ratio is 0.1844 (see Table 1), which amounts to a 40% decrease with respect

to the economy without debt. The ratio of total government expenditures [including the net service of debt] is still higher than the first best—it is almost 23%—, but this still represents a considerably lower weight of government spending in the economy. Likewise, private consumption and capital are brought up closer to the efficient allocation. The ability of Markovian governments to issue debt is thus bound to have sizable positive effects on welfare.

3.4 Numerical Approach

In this section we outline our strategy for the computation of the Markov-perfect equilibrium. The first challenge in the computation of the three unknown functions $C(K, B, \tau)$, $\psi_\tau(K, B)$, and $\psi_B(K, B, \tau)$ stems from the presence of the derivatives of the consumption function in the two generalized Euler equations, (3.13) and (3.14). In a steady state, these derivatives must be solved for, thus making the number of unknowns exceed the number of equations.

Our computational method is an application of a projection method which approximates the three unknown functions with a combination of Chebyshev polynomials. Within the class of orthogonal polynomials, Chebyshev polynomials stand out for its efficiency to approximate smooth functions.⁴ The unknown coefficients in the approximate functions are then obtained so that they satisfy the three Euler equations at some collocation points within a subset of the state space, $[K_{min}, K_{max}] \times [B_{min}, B_{max}]$.

Thus, we approximate functions for consumption, taxes and the issue of debt by:

$$\hat{C}(K, B, \tau; \vec{a}) = \sum_{i=0}^{n_k^c} \sum_{j=0}^{n_b^c} \sum_{\ell=0}^{n_\tau^c} a_{ij\ell} \phi_{ij\ell}(K, B, \tau) \quad (3.19)$$

$$\hat{\psi}_\tau(K, B, \vec{d}) = \sum_{i=0}^{n_k^b} \sum_{j=0}^{n_b^b} d_{ij} \phi_{ij}(K, B) \quad (3.20)$$

$$\hat{\psi}_B(K, B, \vec{h}) = \sum_{i=0}^{n_k^\tau} \sum_{j=0}^{n_b^\tau} h_{ij} \phi_{ij}(K, B), \quad (3.21)$$

where $\phi_{ij\ell}(K, B, \tau)$ and $\phi_{ij}(K, B)$ are tensor products of univariate Chebyshev polynomials, which form the multidimensional basis for approximation.⁵ For instance, $\phi_{ij}(K, B) =$

⁴For a complete characterization of their properties and a rigorous exposition of projection techniques see Judd (1992, 1998). For a previous application of these ideas to the computation of Markovian optimal taxes see Ortigueira (2006).

⁵The debt policy function $\psi_B(K, B, \tau)$ is approximated by $\hat{\psi}_B(K, B, \vec{h})$ as τ may be replaced in the former

$\phi_i(K)\phi_j(B)$, with $\phi_i(K)$ denoting the Chebyshev polynomial of order i in K and $\phi_j(B)$ the Chebyshev polynomial of order j in B . Since Chebyshev polynomials are only defined in the interval $[-1, 1]$, K and B must be re-scaled accordingly, using the chosen $K_{\min}, K_{\max}, B_{\min}, B_{\max}$. That is,

$$\phi_{ij}(K, B) = \phi_i\left(\frac{2(K - K_{\min})}{K_{\max} - K_{\min}} - 1\right) \times \phi_j\left(\frac{2(B - B_{\min})}{B_{\max} - B_{\min}} - 1\right). \quad (3.22)$$

Vectors \vec{a} , \vec{d} , \vec{h} in (3.19)–(3.21) are the unknown coefficients, which are pinned down by imposing that $\hat{C}(K, B, \tau; \vec{a})$, $\hat{\psi}_\tau(K, B, \vec{d})$ and $\hat{\psi}_B(K, B, \vec{h})$ satisfy the three Euler equations and the laws of motion at a number of collocation points. The number of collocation points is set so that the number of equations equals the number of unknown coefficients. In our exercise we choose Chebyshev collocation. It should be noted that the approximation of the debt policy, equation (3.21), embeds already the approximation of the tax policy in terms of K and B . On the other hand, the approximation of the consumption function, (3.19), must be done in terms of K , B and τ , in order to obtain the derivatives of the consumption function which show up in the Generalized Euler equations.

The value function, $W(K, B)$, can then be easily computed as follows. Using the solutions for consumption, taxation and the issue of debt, the value function is approximated by,

$$\hat{W}(K, B, \vec{e}) = \sum_{i=0}^{n_k^v} \sum_{j=0}^{n_b^v} e_{ij} \phi_{ij}(K, B), \quad (3.23)$$

where the vector \vec{e} contains the unknown coefficients in the value function, which are pinned down so that (3.23) solves the government's Bellman equation at a number of collocation points.

3.5 Conclusions

This paper analyzes Markov-perfect optimal fiscal policy in a neoclassical economy with physical capital and public debt. We extend a recent literature on time-consistent policies to economies where the government chooses government expenditures and households hold physical capital and public debt in their portfolios. Previous studies on Markov-perfect policy abstract from either public debt, by assuming a government's period-by-period balanced budget constraint, or from physical capital, assuming that labor is the only factor of production.

by $\psi_\tau(K, B)$. This tax policy function is set at the beginning of the period by the same government that chooses $\psi_B(K, B, \tau)$.

We characterize and compute Markov-perfect optimal fiscal policy in our model economy and find two steady-state equilibrium configurations. We prove that the steady-state Ramsey equilibrium is a Markov-perfect equilibrium. In addition, our numerical computations find a stable, steady-state Markov-perfect equilibrium with positive income taxation and positive public debt. In a calibrated version of the model, this latter equilibrium yields an income tax rate close to 20% and a debt-GDP ratio in the order of 60%. These numbers are in line with those observed in most developed economies. Although the framework presented in this paper displays an expectations-driven multiplicity of equilibria—and thus fails to provide predictions on optimal policy—we argue that it can however be useful as a positive theory of fiscal policy. That is, on how actual policies are determined. The equilibrium with no distortions involves initial tax rates and levels of government asset holdings which may not be feasible in most fiscal constitutions. This could leave the equilibrium with positive long-run distortions as the only Markov-perfect equilibrium of our economy.

Our framework is rather stylized. We have abstracted from endogenous labor supply to focus instead on the role of public debt in economies where the government lacks the ability to commit to future policies. Although endogenizing labor supply, and allowing then the government to set different taxes on capital and labor, would not change our results qualitatively, it might have important quantitative implications. However, the computational costs associated with an extension of our framework in that direction are likely to be insurmountable. As it was made clear in Section 3.4, we are approximating three unknown functions—consumption as a function of capital, debt and taxes, and two government policies as functions of capital and debt—to solve three functional equations. Adding two new unknown functions—labor supply and the tax policy on labor—and two new functional equations will impair the application of projection methods in our set-up.

3.6 Appendix

Proof of Proposition 1:

The problem solved by a government with full commitment is to set infinite sequences

$\{G_t, \tau_t, B_t\}$ so that the implied competitive equilibrium maximizes welfare. That is,

$$\begin{aligned} & \max_{\{G_t, \tau_t, B_{t+1}\}} \sum_{t=0}^{\infty} \beta^t U(C_t, G_t) \\ & s.t. \end{aligned} \quad (3.24)$$

$$C_t + K_{t+1} + G_t = f(K_t) + (1 - \delta)K_t \quad (3.25)$$

$$G_t + [1 + r_t - \delta]B_t = B_{t+1} + \tau_t[(r_t - \delta)(K_t + B_t) + \omega_t] \quad (3.26)$$

$$U_c(C_t, G_t) = \beta U_c(C_{t+1}, G_{t+1})[1 + (1 - \tau_{t+1})(r_{t+1} - \delta)], \quad t = 0 \dots \infty, \quad (3.27)$$

K_0 and B_0 are given.

After defining new variables $\tilde{r}_t \equiv (1 - \tau_t)(r_t - \delta)$ and $\tilde{\omega}_t \equiv (1 - \tau_t)\omega_t$, and formulating the problem of the government as choosing after-tax rental prices, the first-order condition with respect to K_{t+1} (by using the primal approach) can be written as,

$$\Gamma_t = \beta [\Lambda_{t+1}(r_{t+1} - \tilde{r}_{t+1})] + \Gamma_{t+1}(1 + r_{t+1} - \delta), \quad (3.28)$$

where Γ_t and Λ_t are Lagrange multipliers. Using the Euler equation, equation (6.5) in a steady-state equilibrium is,

$$(\Gamma + \Lambda)(r - \tilde{r}) = 0, \quad (3.29)$$

from which it follows that $\tau = 0$ in the steady-state equilibrium, and, consequently, $B < 0$.

Proof of Proposition 2:

The first-order condition to B' in government's maximization problem (3.11) is given by,

$$U_G G_{B'} - \beta \tilde{V}'_{K'} G_{B'} + \beta \tilde{V}'_{B'} = 0. \quad (3.30)$$

The first-order condition to τ in government's maximization (3.12) is,

$$U_c C_\tau + U_G (G_\tau + G_{B'} B'_\tau) - \beta \tilde{V}'_{K'} (C_\tau + G_\tau + G_{B'} B'_\tau) + \beta \tilde{V}'_{B'} B'_\tau = 0, \quad (3.31)$$

which, after making use of (3.30), simplifies to,

$$U_c C_\tau + U_G G_\tau - \beta \tilde{V}'_{K'} (C_\tau + G_\tau) = 0. \quad (3.32)$$

Envelope conditions, along with $W(K, B) = \tilde{V}(K, B)$, yield,

$$W_K = U_c C_K + U_G G_K + \beta W'_{K'} [1 + f_K - \delta - C_K - G_K] \quad (3.33)$$

$$W_B = U_c C_B + U_G G_B + \beta W'_{K'} [-C_B - G_B]. \quad (3.34)$$

Forwarding these envelope conditions one period and using the above first-order conditions, we obtain the two Generalized Euler Equations, (3.13) and (3.14), presented in Proposition 2.

Proof of Proposition 3:

As shown in Proposition 1, in a steady-state Ramsey equilibrium income taxes are zero and the government holds negative debt (assets) to finance the provision of the public good. The government does not rely on distortionary taxation, and the efficiency condition, $U_c = U_G$, is attained. In this proof we show that the system of equations characterizing steady-state Markov-perfect equilibria has a solution with these properties.

Let us start by assuming that $U_c = U_G$. Then, from (3.32) it follows that $U_c = \beta W_K$. From (3.34) it is then easy to see that $W_B = 0$. Finally, equation (3.33) becomes,

$$\frac{1}{\beta} = 1 + f_K - \delta, \quad (3.35)$$

which, along with the consumer's Euler equation, implies that $\tau = 0$.

Proof of Proposition 4:

The proof follows directly from the first-order and envelope conditions presented above, along with the non-arbitrage condition. Thus, by plugging the first-order condition to issues of debt evaluated at a steady-state Markov-perfect equilibrium into (3.34), it obtains that at a steady-state equilibrium,

$$(1 + \beta G_B)W_B = (U_c - U_G - \beta W_B)C_B. \quad (3.36)$$

Then, plugging $G_B = -(1 - \tau)q - 1$ and the non-arbitrage condition, $q = f_K - \delta$, into equation (3.36) and using the household's Euler equation, it follows that $C_B = 0$ at the steady state.

Proof of Proposition 5:

Here we prove that the two steady-state equilibria—one with positive distortions and one without—are not associated with the same pair of decision rules ψ_τ and ψ_B . To do this, we show that the policy rules generating the steady state with no distortions are the only limit of policy rules in the finite-horizon economy as the planning horizon goes to infinity. The proof, although algebraically tedious, is straightforward.

In the finite-horizon economy with last period denoted by T , we have $K_{T+1} = B_{T+1} = 0$. Therefore, in period T households simply consume all their resources. The problem of the

time- T government is then,

$$\max_{\tau_T} U(C_T, G_T)$$

s.t.

$$C_T = K_T + B_T + (1 - \tau_T) [f(K_T) - \delta K_T + q_T B_T] \quad (3.37)$$

$$G_T = \tau_T [f(K_T) - \delta K_T + q_T B_T] - (1 + q_T) B_T. \quad (3.38)$$

The first-order condition to this problem is,

$$U_c(T) = U_G(T), \quad (3.39)$$

where $U_c(T)$ denotes $U_c(C_T, G_T)$.

In period $T - 1$, the households' Euler equation is,

$$U_c(C_{T-1}, G_{T-1}) = \beta U_c(C_T, G_T) [1 + (1 - \tau_T) (f_K(K_T) - \delta)], \quad (3.40)$$

and the non-arbitrage condition between the two assets is $q_T = f_K(K_T) - \delta$. The fiscal policy chosen by the time- $(T-1)$ government is obtained in a two-step maximization problem. First, given τ_{T-1} , the issue of debt solves,

$$\max_{B_T} \{U(C_{T-1}, G_{T-1}) + \beta U(C_T, g_T)\}$$

s.t.

$$G_{T-1} = B_T + \tau_{T-1} [f(K_{T-1}) - \delta + q_{T-1} B_{T-1}] - (1 + q_{T-1}) B_{T-1} \quad (3.41)$$

$$K_T = f(K_{T-1}) + (1 - \delta) K_{T-1} - G_{T-1} - C_{T-1} \quad (3.42)$$

and equations (3.37), (3.38), (3.39) and (3.40).

The first-order condition to this problem is,

$$U_G(T-1) \frac{\partial G_{T-1}}{\partial B_T} + \beta \left[\left(U_c(T) \frac{dC_T}{dK_T} + U_G(T) \frac{dG_T}{dK_T} \right) \frac{dK_T}{dB_T} + \left(U_c(T) \frac{dC_T}{dB_T} + U_G(T) \frac{dG_T}{dB_T} \right) \right] = 0 \quad (3.43)$$

Then, τ_{T-1} is the solution to,

$$\max_{\tau_{T-1}} \{U(C_{T-1}, G_{T-1}) + \beta U(C_T, g_T)\}$$

s.t.

equations (3.37), (3.38), (3.39), (3.40), (3.41), (3.42) and (3.43).

The first-order condition is,

$$U_c(T-1) \frac{\partial C_{T-1}}{\partial \tau_{T-1}} + U_G(T-1) \frac{\partial G_{T-1}}{\partial \tau_{T-1}} + \beta \left[U_c(T) \frac{dC_T}{dK_T} + U_G(T) \frac{dG_T}{dK_T} \right] \frac{\partial K_T}{\partial \tau_{T-1}} = 0 \quad (3.44)$$

Now, from feasibility conditions at T and $T - 1$ we obtain that $\frac{dC_T}{dB_T} = -\frac{dG_T}{dB_T}$, $\frac{\partial K_T}{\partial \tau_{T-1}} = -\frac{\partial C_{T-1}}{\partial \tau_{T-1}} - \frac{\partial G_{T-1}}{\partial \tau_{T-1}}$ and $\frac{dK_T}{dB_T} = -\frac{dG_{T-1}}{dB_T} = -1$. Using these values in equations (3.43) and (3.44), we have

$$[U_c(T-1) - U_G(T-1)] \frac{\partial C_{T-1}}{\partial \tau_{T-1}} = 0. \quad (3.45)$$

Since $\frac{\partial C_{T-1}}{\partial \tau_{T-1}} \neq 0$, it thus follows that $U_c(T-1) = U_G(T-1)$. Then, using the fact that $\frac{dC_T}{dK_T} + \frac{dG_T}{dK_T} = 1 + f_K(K_T) - \delta$, equation (3.44) yields,

$$U_c(T-1) = \beta U_c(T) [1 + f_K(K_T) - \delta]. \quad (3.46)$$

From this equation and the household's Euler equation it follows that $\tau_T = 0$.

Solving the problem for period $T - 2$, yields $\tau_{T-1} = 0$. By proceeding in this way up to the initial period, it can be shown that all taxes are zero but the initial one. That is, $\tau_0 \neq 0$ and $\tau_t = 0$ for all t from 1 to T .

Figure 1. Debt Policy Function

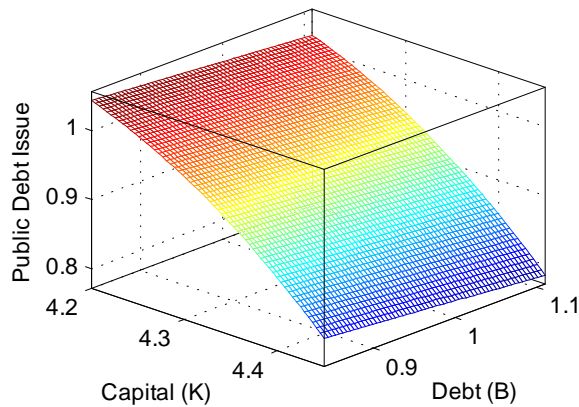


Figure 2. Tax Policy Function

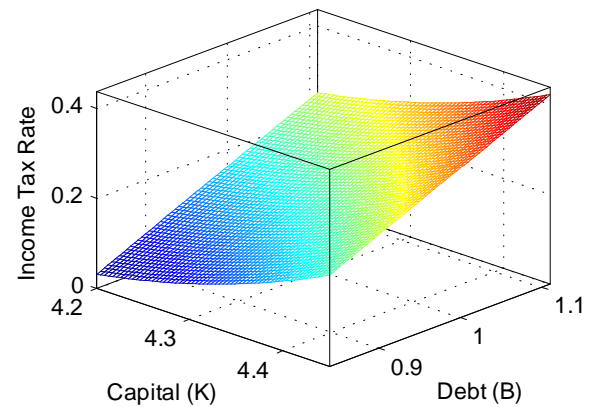


Figure 3. Consumption Function

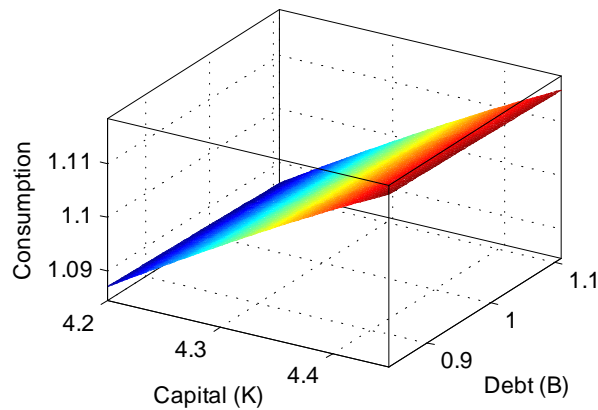
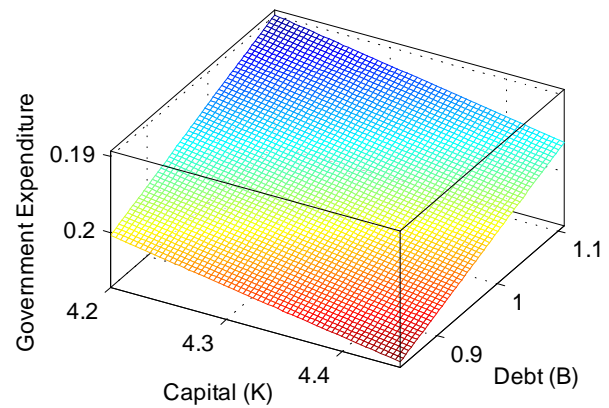


Figure 4. Government Expenditure Function



Notes: Figures 1 to 4 show the policy functions in a Markov-perfect equilibrium. The government's tax policy is shown in Figure 1. The government's debt policy is shown in Figure 2. The government's spending policy is displayed in Figure 3. Finally, the private consumption function is shown in Figure 4.

Figure 5. Net Investment in Physical Capital

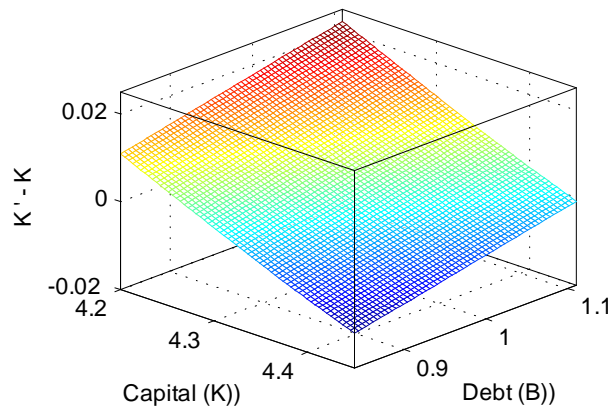
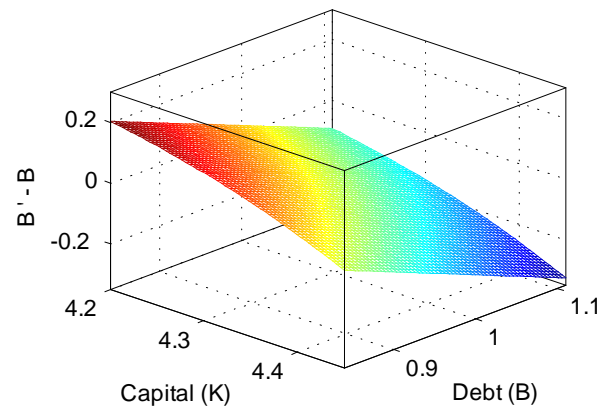
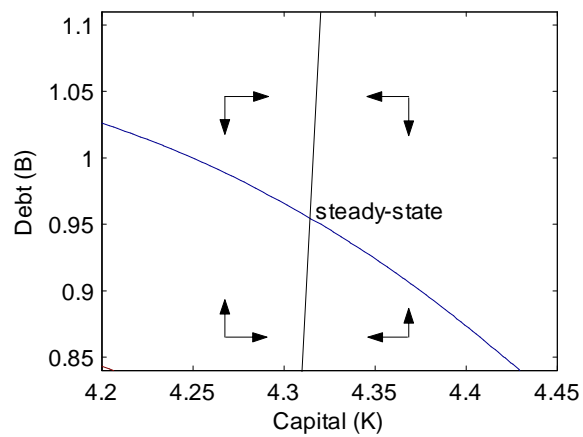
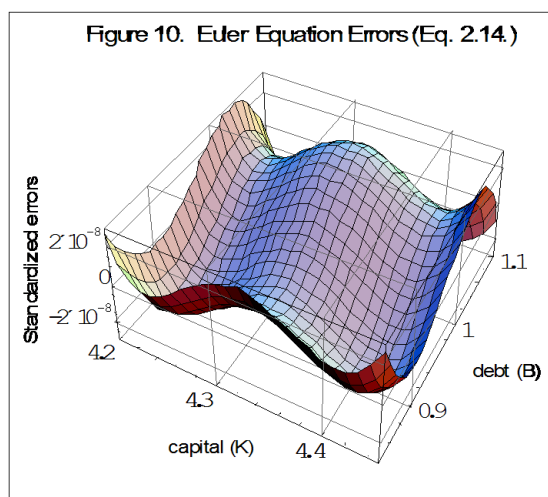
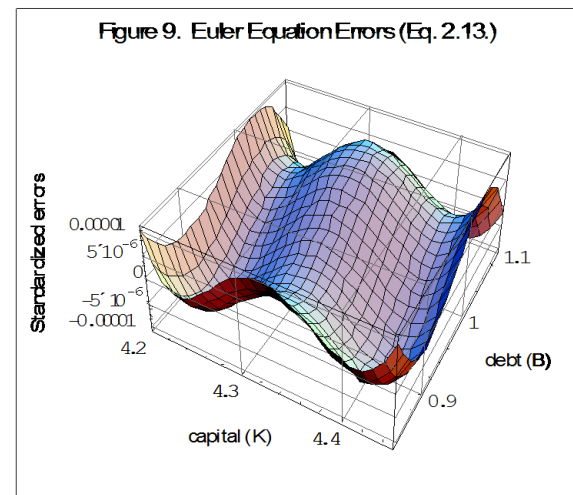
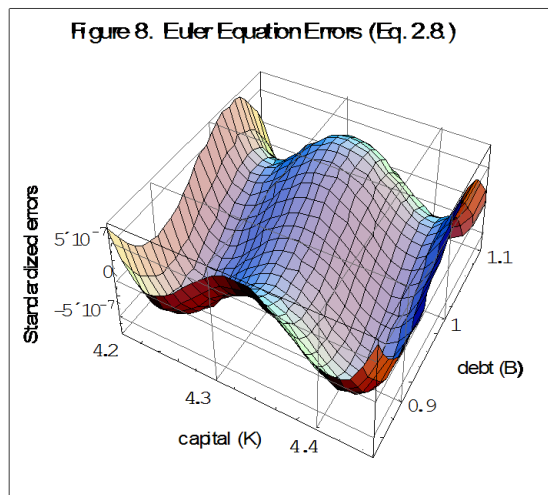


Figure 6. Issue of Public Debt

Figure 7. The $K' = K$ and $B' = B$ Loci

Notes: Figures 5 to 7 show the dynamics around the steady-state equilibrium with positive income taxation. Figure 5 shows net investment; Figure 6 shows the change in government debt; and Figure 7 shows the $K' = K$ and $B' = B$ loci.



Notes: Figures 8 to 11 show relative errors of Chebyshev collocation for the Euler equation (Figure 8), Generalized Euler equations (Figures 9 and 10) and the Bellman equation (Figure 11).

3.7 References

- Aiyagari, R., Marcet, A., Sargent, T., Seppälä (2002). "Optimal Taxation without State-Contingent Debt." *Journal of Political Economy*, vol. 110, no. 6, pp. 1220–1254.
- Alesina, A., and Tabellini, G., (1990). "A Positive Theory of Fiscal Deficits and Government Debt." *Review of Economic Studies*, vol. 57, no. 3, pp. 403–414.
- Azzimonti-Renzo, M., Sarte, P. D., and Soares, J., (2006). "Optimal Policy and (the lack of) time inconsistency: Insights from simple models." Manuscript.
- Calvo, G., (1988). "Servicing the Public Debt: The Role of Expectations." *American Economic Review*, vol. 78, no. 4, pp. 647–661.
- Chamley, C., (1986). "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives." *Econometrica*, vol. 54, no.3, pp. 607–622.
- Debortoli, D. and Nunes, R., (2007). "Political Disagreement, Lack of Commitment and the Level of Debt." Manuscript.
- Díaz-Giménez, J., Giovannetti, G., Marimon, R. and Teles, P. (2007). "Nominal Debt as a Burden on Monetary Policy." *Review of Economic Dynamics*, *forthcoming*.
- Judd, K., (1992). "Projection Methods for Solving Aggregate Growth Models." *Journal of Economic Theory*, vol. 58, pp. 410–452.
- Judd, K., (1985). "Redistributive Taxation in a Simple Perfect Foresight Model." *Journal of Public Economics*, vol. 28, pp. 59–83.
- Judd, K., (1998). *Numerical Methods in Economics*. Cambridge, Mass.: MIT Press.
- Klein, P. and Ríos-Rull, J., (2003). "Time-Consistent Optimal Fiscal Policy." *International Economic Review*, vol. 44, no. 4, pp. 1207–1405.
- Klein, P., Krusell, P., and Ríos-Rull, J.-V., (2006). "Time-Consistent Public Policy." *Review of Economic Studies*, *forthcoming*.
- Krusell, P., Martin, F., and Ríos-Rull, J., (2006). "Time-Consistent Debt." Manuscript.
- Lucas, R. and Stokey, N., (1983). "Optimal Fiscal and Monetary Policy in an Economy without Capital." *Journal of Monetary Economics*, vol. 12, no. 1, pp. 55–94.
- Martin, F., (2006). "A Positive Theory of Government Debt." Manuscript.
- Ortigueira, S., (2006). "Markov-perfect optimal Taxation." *Review of Economic Dynamics*, vol. 9, pp. 153–178.
- Persson, M., Persson, T. and Svensson, L., (1988). "Time Consistency of Fiscal and Monetary Policy." *Econometrica*, vol. 55, pp. 1419–1432.

Persson, T. and Svensson, L., (1989). "Why a Stubborn Conservative would Run a Deficit: Policy with Time-Inconsistent Preferences." *The Quarterly Journal of Economics*, vol. 104, no. 2, pp. 325–345.

Song, Z., Storesletten, K. and Zilibotti, F., (2007). "Rotten Parents and Disciplined Children: A Politico-Economic Theory of Public Expenditure and Debt." Manuscript.